

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:32:52 ; Search time 168.5 Seconds
(without alignments)
4483.584 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atgcggtcccggtttgggaa.....cagctcagcaagaagcttaa 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO spool_p/US10089688/runat 10112004 152246 4835/app query.fasta_1.1223
-DB=A Geneseq 23Sep04 -OPMT=fastan -SUFFIX=rag -MINMATCH=0 1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THRE SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089688 @CGN 1 1 224 @runat 10112004 152246 4835 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1781	98.8	350	AAB20387	Aab20387 Human acu
2	1635.5	90.7	341	AAB48970	Aab48970 Human ANI
3	1635.5	90.7	341	ADJ69105	Adj69105 Human hea
4	1635.5	90.7	496	AAB10858	Aae10858 Gal4-huma
5	1635.5	90.7	552	AAB10859	Aae10859 LexA-huma
6	1614.5	89.5	341	AAY94248	Aay94248 Mouse cal
7	1597.5	88.6	354	ABG23844	Abg23844 Novel hum
8	1297.5	72.0	337	AAY94247	Aay94247 Human cal
9	1297.5	72.0	337	AAB82090	Aab82090 Human acu
10	1297.5	72.0	337	AAM39078	Aam39078 Human pol

11	1153.5	64.0	289	4	AAB94139	Aab94139 Human pro
12	1123	62.3	339	4	ABB60392	Abb60392 Drosophil
13	1117	62.0	339	3	AAI94249	Aai94249 Drosophil
14	1054.5	58.5	377	3	AAI94250	Aai94250 C. elegan
15	888.5	49.3	237	4	AAM40864	Aam40864 Human pol
16	705.5	39.1	343	3	AAG45273	Aag45273 Arabidops
17	674	37.4	326	3	AAG51053	Aag51053 Arabidops
18	674	37.4	400	3	AAG51052	Aag51052 Arabidops
19	674	37.4	504	3	AAG51051	Aag51051 Arabidops
20	673.5	37.4	300	3	AAG23886	Aag23886 Arabidops
21	669.5	37.1	345	3	AAG05089	Aag05089 Arabidops
22	666.5	37.0	300	3	AAG45274	Aag45274 Arabidops
23	666.5	37.0	300	3	AAG30714	Aag30714 Arabidops
24	666.5	37.0	305	3	AAG30713	Aag30713 Arabidops
25	658.5	36.5	290	3	AAG23887	Aag23887 Arabidops
26	635	35.2	320	3	AAG05090	Aag05090 Arabidops
27	523.5	29.0	213	3	AAG23888	Aag23888 Arabidops
28	518.5	28.8	213	3	AAG30715	Aag30715 Arabidops
29	518.5	28.8	213	3	AAG45275	Aag45275 Arabidops
30	465.5	25.8	154	3	AAG41151	Aag41151 Zea mays
31	460	25.5	197	3	AAG05091	Aag05091 Arabidops
32	448.5	24.9	148	3	AAG41152	Aag41152 Zea mays
33	430.5	23.9	139	3	AAG41153	Aag41153 Zea mays
34	276	15.3	639	4	ABG25372	Abg25372 Novel hum
35	271	15.0	383	4	ABG23843	Abg23843 Novel hum
36	269	14.9	236	5	ABP02921	Abp02921 Human ORF
37	209	11.6	135	5	ABP34081	Abp34081 Human ORF
38	173	9.6	75	4	ABG25370	Abg25370 Novel hum
39	127.5	7.1	121	3	AAG33914	Aag33914 Arabidops
40	126	7.0	1182	6	ABU25521	Abu25521 Protein e
41	116.5	6.5	888	5	AAB3013	Aau3013 Human hom
42	116.5	6.5	972	6	ABU19394	Abj19394 Human int
43	113	6.3	541	7	ADH87535	Adh87535 Enterococ
44	113	6.3	2184	4	AAE00425	Aae00425 P. falcip
45	111.5	6.2	790	7	ADJ70372	Adj70372 Human hea

ALIGNMENTS

RESULT 1
AAB20387

ID AAB20387 standard; protein; 350 AA.

XX AAB20387;

XX 11-JUN-2001 (first entry)

DE Human acute neuronal induced calcium binding protein ANIC-BP-1B.

XX Acute neuronal induced calcium binding protein; ANIC-BP-1B;
KW spice variant; human; stroke; head trauma; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; spinal cord injury;
KW cerebroprotective; antiparkinsonian; nootropic; neuroprotective; therapy;
KW diagnosis; vaccine.

XX Homo sapiens.

XX WO200125423-A1.

PD 12-APR-2001.

XX 28-SEP-2000; 2000WO-EP009475.

XX 04-OCT-1999; 99EP-00119113.

XX (MERE) MERCK PATENT GMBH.

XX Duecker K, Den Daas I;

XX WPI; 2001-266306/27.

XX N-PSDB; AAF30688.

PT Novel human acute neuronal induced calcium-binding protein like protein

PT splice variant, useful for treating stroke, acute head trauma,
PT Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal cord
PT injury.
XX
XX
PS
PS
XX

Claim 2; Page 44-45; 49pp; English.

CC The present sequence is that of a novel human acute neuronal induced
CC calcium binding protein-like protein splice variant, ANIC-BP-1B. The
CC protein shows homology to other members of the calcium binding protein
CC family, including ANIC-BP, a protein discovered by mRNA differential
CC display that is upregulated in a rat model of head trauma. ANIC-BP and
CC ANIC-BP-1B differ in their C-terminal portions. The variant protein could
CC serve as a novel drug target. The invention provides ANIC-BP-1B
CC polynucleotides (see AAF30688) and polypeptides, expression vectors, host
CC cells and antibodies, as well as methods for producing the protein and
CC for treating or preventing disorders associated with expression of the
CC protein by inhibiting or activating the action of ANIC-BP-1B. Diseases
CC that may be treated include stroke and acute head trauma, Parkinson's
CC disease, Alzheimer's disease, multiple sclerosis and spinal cord injury.
CC The polynucleotides and polypeptides can also be used in diagnostic
CC assays and in vaccines, and to identify agonists and antagonists useful
CC for treating conditions associated with ANIC-BP-1B imbalance

XX
SQ Sequence 350 AA;

Alignment Scores:
Pred. No.: 2.53e-178 Length: 350
Score: 1781.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.78% Indels: 0
DB: 4 Gaps: 0

US-10-089-688-1 (1-1053) x AAB20387 (1-350)

Qy 1 ATGCGGTCCCGTTGGAGCTCTCAAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
Db 1 MetProPheProPheGlySerHisLysSerProAlaAspIleValLysAsnLeuLys 20
Qy 61 GAGAGCTAGGCTGTTCTGGAAGAAGCAACATTTCTGATATAAAAGCAGAAAGGCTACA 120
Db 21 GluSerMetAlaValLeuLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
Qy 121 GAAGAAGTTTCAAAATCTGGTCCCATGAAAGAAATCTGTATGGCACAATATGAAAAA 180
Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
Qy 181 GAGCTCAGACAGACAGTAGCTCACTCTCTCAAGAACTCTATATAGTGGCTCCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
Qy 241 AGCACCTCTGGTAGCTGATTATACAGCTCAATGACTTTTGAGGGCAAAAAGAGCTGGCTCAA 300
Db 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
Qy 301 ATTTTCAACAATATCTCAGAGAACAATTTGGTAGCAGAACTCTCTACTGTTGAATACATC 360
Db 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
Qy 361 TGCACCCCAACAGAAATTTTGTCTATGTTTAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
Qy 421 CTAATTTGGAATAATGTTAAGAATGTCATCAGACATGACCACTTGCAGAAATCATTT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
Qy 481 TTGTGGTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCGAAATTTGACATTA 540
Db 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
Qy 541 GCTTCAGATGCTATTTGCCACATTTCAAGGATTTTACTTACAAGACATAAATTTGCTCAGTGA 600

Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
Qy 601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATCAGAGTACTTCAATTC 660
Db 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
Qy 661 GAAAAATTATGTGACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
Qy 721 CACAACCTTCAAAATATGACAAATATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
Qy 781 AACTGCTGGCAGACAAAAGTCCCAACATCCAGTTTGAGGCTTTCACGTTTTTAAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTTGTAGCCAATCTTAACAGACGCGAGCCCATCTAGACATCTCTCTCAAGAACCCAGGCC 900
Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLysAsnGlnAla 300
Qy 901 AAACTCATAGATTTCTCAGCAAGTTTTCAGAACGACAGGACGAGTGTATGAGCAGTTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrAspCysMetSerSerSer 320
Qy 961 GTACCGAGCAGAAATTTCCCGGGTCGATTACGGGTTAAACCGCGGAGCGTGGATCAGG 1020
Db 321 ValProThrThrAsnSerArgValAspLeuArgValLysProArgThrArgGlyIleArg 340
Qy 1021 GATTTGAAGACACAGCTCAGCAAGAGCT 1050
Db 341 AspLeuLysArgProAlaGlnGlnGluAla 350

RESULT 2

AAB48970

ID AAB48970 standard; protein; 341 AA.

XX AAB48970;

XX 27-MAR-2001 (first entry)

XX Human ANIC-BP (acute neuronal induced calcium-binding protein).

XX Human; acute neuronal induced calcium-binding protein; ANIC-BP;

XX Mo25 homologue; Hyma homologue; drug screening; stroke;

XX acute head trauma; multiple sclerosis; spinal cord injury; vaccine;

XX cerebroprotective; neuroprotective.

XX Homo sapiens.

XX WO200078947-A1.

XX 28-DEC-2000.

XX 14-JUN-2000; 2000WO-EP005457.

XX 22-JUN-1999; 99EP-00112024.

XX (MERE) MERCK PATENT GMBH.

XX Den Daas I, Fischer V, Seyfried C, Von Melchner L;

XX WPI; 2001-102721/11.

XX N-PSDB; AAC91772.

XX Novel acute neuronal induced calcium binding protein, useful for treating

XX acute head trauma, stroke, multiple sclerosis and spinal cord injury.

XX Claim 2; Page 37; 50pp; English.

XX The invention relates to human acute neuronal induced calcium-binding

XX protein (ANIC-BP) and to nucleic acid encoding it. The invention also

CC relates to expression systems and recombinant host cells comprising ANIC-
 CC BP DNA, the recombinant production of ANIC-BP, antibodies specific for
 CC ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin Fc
 CC region, and methods of screening for modulators of ANIC-BP function. ANIC
 CC -BP has homology and structural similarity to HymA and Mo25 proteins.
 CC ANIC-BP proteins and nucleotides are useful for treating stroke and acute
 CC head trauma, multiple sclerosis and spinal cord injury. ANIC-BP proteins
 CC are useful in screening assays, for identifying membrane bound or soluble
 CC receptors, and also in vaccines. ANIC-BP nucleotides are useful as
 CC diagnostic reagents, as tools for tissue expression studies, for
 CC chromosome localisation studies, as genetic vaccines, and in the
 CC generation of transgenic animals. The present sequence represents human
 CC ANIC-BP
 XX
 SQ Sequence 341 AA;

Alignment Scores:
 Pred. No.: 5,43e-163 Length: 341
 Score: 1635.50 Matches: 329
 Percent Similarity: 94.29% Conservative: 1
 Best Local Similarity: 94.00% Mismatches: 11
 Query Match: 90.71% Indels: 9
 DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AAB48970 (1-341)

QY 1 ATCCGCTCCGTTGGAGTCTCAAAATCTCCAGACAGATTTGAAGAATCTGAAG 60
 DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
 QY 61 GAGAGCATGCTGTTCTGGAAGCAGACAGATTTCTGATTAAGAAGCAGAAAGGCTACA 120
 DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAAGTTTCCAAAATCTGGTTGCCATCAAGAAATTTCTGTATGGCACAATGAAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
 QY 181 GAGCTCAGACAGACAGTAGTCTAACTTGCTCAAGAACTCTATATATGAGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuAlaGlnLeuLysSerGlyLeuLeu 80
 QY 241 AGCACCTGCTAGTCTGATTTACAGCTCATTTGACCTTTGAGGGCAAAAAGACGTTGCTCAA 300
 DB 81 SerThrLeuValAlaAspLeuGlnLeuLeuAlaAspPheGluGlyLysLysAspValAlaGln 100
 QY 301 ATTTTCAACAATATTTCTCAGACAGACAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
 DB 101 IlePheAsnAsnIleLeuAArgGlnIleGlyThrArgThrProThrValGluLysIle 120
 QY 361 TGCACCCAAACAGAAATATTTGTTTCATGTTATTTGAAGGGTATGAATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
 QY 421 CTRAAATGTCGAATATGTTAAAGATGTCATCAGACATCAACCACTTGCAAAAATCAAT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
 QY 481 TTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACAAATTTGACATA 540
 DB 161 LeuTrpSerGluGlnPheThrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATGATTTGCCAATCAAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
 DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 QY 601 GAATTTTGCAGACAGATTTATGATAGATTTTTCAGTGAATATGAGAGTACTTCAATCA 660
 DB 201 GluPheLeuGluGlnHisLysAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 QY 661 GAAATTTATGTGACAAAAGACAGTCACTGAAGCTTCTCGGTCAACTACTACTAGATAGA 720
 DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240

QY 721 CACAACTTCAATATTATGACAAATATACATCAGTAAACCTGAGAACCTCAAATTAATCATG 780
 DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 QY 781 AACTCTGCTGAGACAAACAACTCCCAACATCCAGCTTTTCAGGCTTTTCAGCTTTTAAGGTG 840
 DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 QY 841 TTTGTAGCCAATCTTAACAGAGCAGCCCATCTCAGACATCTCTCCTCAAGAACCAAGGCC 900
 DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
 QY 901 AAACCTCATAGTCTCTCAGCAAGTTTTCAGAACGACAGAGCGAGATTGTATGAGCAGTTCC 960
 DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 QY 961 GTACCCGAGCAGCAATTCCTCCGGTTCGATTACGGCTTAAACCGCGGACGCTGGGATCAGG 1020
 DB 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 QY 1021 GATTTCAGAGACAGCAGCTCAGCAAGAGCT 1050
 DB 332 AspLeuLysArgProAlaGlnGlnGluAla 341

RESULT 3
 ADJ69105
 ID ADJ69105 standard; protein; 341 AA.
 XX
 AC ADJ69105;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SeqID911.
 XX
 KW mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 PR 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-0389987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 PA (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 XX
 PI Chosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
 PI Warnock DE;
 XX
 DR WPI; 2003-845369/78.
 XX
 PT Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX
 PS Claim 1; SEQ ID NO 911; 180pp; English.
 CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for

identifying proteins of the human heart mitochondrial proteome that are useful for drug screening assays, as well as therapeutic targets. The present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, nootropic, antidiabetic, anticonvulsant, antiarrhythmic, osteopathic, ophthalmological and cytostatic activities. This polypeptide sequence is a human heart mitochondrial protein of the invention.	
XX	Sequence 341 AA;
SQ	
Alignment Scores: 5 43e-163 Length: 341	
Pred. No.:	1635.50 Matches: 329
Score:	94.29% Conservative: 1
Best Local Similarity:	94.00% Mismatches: 11
Query Match:	90.71% Indels: 9
DB:	7 Gaps: 2
US-10-089-688-1 (1-1053) x ADJ69105 (1-341)	
QY	1 ATGCGGTTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
DB	1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY	61 GAGAGCATGCTGTTCTGGAAAGCAGACATTTCTGATTAATAAAAGCAGAAAAGGCTACA 120
DB	21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY	121 GAAGAAGTTTCCAAAATCTGGTGGCCATCAAGAAATTTCTGTATGGCACAATGAAAAA 180
DB	41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuTyrGlyThrAsnGluLys 60
QY	181 GAGCCTCAGACAGACAGTAGCTCAACTGCTCAAGAACTCTATATATAGTGGCTCCTT 240
DB	61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
QY	241 AGCACCTGTGTAGTGTATACACTCATTGACCTTTCAGGGCAAAAAGAGCTGGCTCAA 300
DB	81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY	301 ATTTTCAACATATTTCTCAGAAACAAATTTGGTACGAGAACTCCTACTGTTTGAATACATC 360
DB	101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY	361 TGCACCCCAAGAAATATTTTGTGTTATTTGAAAGGTATGAATCTCCAGAAATAGCT 420
DB	121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
QY	421 CTAAATTTGGAATATGTTAAGAAATGCATCAGACATGAACCACTTGCAGAAATCATT 480
DB	141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY	481 TTGTGCTCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCACAAATTTGACATA 540
DB	161 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
QY	541 GCTTCAGATGCAATTTGCCACATTCAGGATTTACTACAAGACATAAATTCCTCAGTGCA 600
DB	181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY	601 GAATTTTGGACACATATGATAGATTTTTCAGTGAATATGAGAAGTACTTCTCATTTCA 660
DB	201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
QY	661 GAAATATTTGTGACAAAAGACAGTCACTGAAGCTTCTCGGTGAAGTACTACTAGATAGA 720
DB	221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 240
QY	721 CACAACTTCACAATATGACAAAATACATCAGTAAACCTGAGAACTCAAAATTAATGATG 780
DB	241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY	781 AACCTGCTCGGAGACAAAGTCGCAACATCCAGTTTGGCCCTTTTCAGCTTTTAAAGGTG 840
DB	261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY	841 TTTGTAGCCATCTTAAACAAGACGCGCCATCTCTAGACATCTCTCAAGAACCCAGGCG 900
DB	281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
QY	901 AAACATCATAGACTTCTCTCAGCAAGTTTTCAGAACGACGAGATTTGTATGACAGTTCC 960
DB	301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu 315
QY	961 GTACCGACGACGAATTCCTCCGGTTCGATTACCGTTAAACCGCGGAGCGTGGGATCAGG 1020
DB	316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln 11eArg 331
QY	1021 GATTTGAAGACGACCTCAGCAAGAGCT 1050
DB	332 AspLeuLysArgProAlaGlnGlnGluAla 341
RESULT 4	
AAE10858	
ID	AAE10858 standard; protein; 496 AA.
XX	AAE10858;
XX	18-DEC-2001 (first entry)
DT	
XX	Gal4-human ANIC-BP-1 fusion protein.
DE	
XX	Human; acute neuronal induced calcium binding protein type 1 ligand;
KW	ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
KW	Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
KW	gene therapy; fusion protein; Gal4 protein.
XX	Homo sapiens.
OS	Unidentified.
OS	Chimeric.
XX	WO200170771-A2.
PN	
XX	27-SEP-2001.
PD	
XX	20-MAR-2001; 2001WO-EP003149.
PF	
XX	21-MAR-2000; 2000EP-00106110.
PR	
XX	(MERE) MERCK PATENT GMBH.
FA	
XX	Den Daas I, Duecker K, Hock B;
PI	
XX	WPI; 2001-607519/69.
DR	
XX	Novel acute neuronal induced calcium binding protein type 1 ligand
PT	polypeptides, useful in the treatment of stroke, head trauma, multiple
PT	sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
PT	injury.
XX	Disclosure; Page 42-44; 46pp; English.
PS	
XX	The invention relates to human acute neuronal induced calcium binding
CC	protein type 1 (ANIC-BP-1) ligand polypeptides and polymucleotides.
CC	Sequences of the invention are useful for treating human diseases
CC	including stroke, head trauma, multiple sclerosis, Parkinson's disease,
CC	Alzheimer's disease and spinal cord injury. They are also useful as
CC	vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
CC	soluble receptors. Polynucleotides of the invention are useful as
CC	diagnostic reagents, for chromosome localization studies, and as valuable
CC	tools for tissue expression studies. They are also useful in gene

CC therapy. The present sequence is Gal4-human ANIC-BP-1 fusion protein
CC comprising the Gal4 protein and a C-terminally linked human ANIC-BP-1
CC protein

SQ Sequence 496 AA;

Alignment Scores:

Pred. No.: 6.26e-163 Length: 496
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AAE10858 (1-496)

QY 1 ATGCGGTTCCCGTTGGGAGCTCTACAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
DB |||||
DB 156 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 175
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAAGCAGAAAGCGCTACA 120
DB |||||
DB 176 GluSerMetAlaValLeuGluLysGluAspIleSerAspLysAlaGluLysAlaThr 195
QY 121 GAAGAAGTTTCCAAATAATCTGGTCCGCTGAAAGAAATCTGTATGGCACAATGAAAA 180
DB |||||
DB 196 GluGluValSerLysAsnLeuValAlaMetLysGluLysGluLysGlyThrAsnGluLys 215
QY 181 GAGCCTCAGACAGAGCAGTAGCTCAACTCTCTCAAGAACTCTAATAATAGTGGCTCCTT 240
DB |||||
DB 216 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysLysSerGlyLeuLeu 235
QY 241 AGCACCCCTGGTAGCTGATTACAGCTCATTCAGCTTTGAGGGCAAAAGAGCGTGGCTCAA 300
DB |||||
DB 236 SerThrLeuValAlaAspLeuGlnLeuLysAspPheGluGlyLysAspValAlaGln 255
QY 301 ATTTTCAACAATATCTCAGAAAGCAAAATGGTAGAGAACTCTACTGTGTAATACATC 360
DB |||||
DB 256 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluThrIle 275
QY 361 TGACCCCAACAGATATTTTCTCATGTTATTGAAGGATGATGAATCTCCAGAAATAGCT 420
DB |||||
DB 276 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyThrGluSerProGluIleAla 295
QY 421 CTAAATTTGTGAATAATGTTAAGAGAATGCAATCAGACATGAAACCACTTGCAGAAATCAT 480
DB |||||
DB 296 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 315
QY 481 TTGTGTCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCACACATTTGACATA 540
DB |||||
DB 316 LeuTrpSerGluGlnPheThrAspPhePheArgTyrValGluMetSerThrPheAspIle 335
QY 541 GCTTCAGATGATTTGCAACATTCAGATTTACTTACAGACATTAATGCTCAGTGCA 600
DB |||||
DB 336 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 355
QY 601 GAAATTTTGGAAACAGCATTTATGATGATTTTTCAGTGAATATGAGAAGTACTTCAATCA 660
DB |||||
DB 356 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluThrGluLysLeuLeuHisSer 375
QY 661 GAAATATGATGACAAAGACAGTCACTAGCTTTCGGTGAACACTACTACTAGATAGA 720
DB |||||
DB 376 GluAsnThrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 395
QY 721 CACAACTTCAATTTGACAAATATCATCATCACTGAACTGAGACCTCAATTAATCATG 780
DB |||||
DB 396 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 415
QY 781 AACCTGCTGCGAGACAAAGTCGCAACATCCAGTTTTCAGGCTTTCAGCTTTTAAAGTG 840
DB |||||
DB 416 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 435
QY 841 TTTGTAGCCAAATCTTAACAGACGACGCCCATCTAGACATCTCTCTCAAGAACCGGCC 900

Db |||||
436 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 455
QY 901 AAATCATAGAGTTCTCTCAGCAAGTTTTCAGAACACAGCAGCGGATTTGTATGAGCAGTTCC 960
DB |||||
456 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 470
QY 961 GTACCGACGACGAATTTCCCGGTCGATTACGGCTTTAAACCGCGGACGCGTGGGATCAGG 1020
DB |||||
471 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 486
QY 1021 GATTGGAAGACACAGCTCAGCAAGAGCT 1050
DB |||||
487 AspLeuLysArgProAlaGlnGlnGluAla 496
RESULT 5
AAE10859
ID AAE10859 standard; protein; 552 AA.
XX
AC AAE10859;
XX
DT 18-DEC-2001 (first entry)
XX
XX LexA-human ANIC-BP-1 fusion protein.
XX
KW Human; acute neuronal induced calcium binding protein type 1 ligand;
KW ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
KW Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
KW gene therapy; fusion protein; LexA protein.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
XX
XX Key Location/Qualifiers
XX Region 1..202
XX /note= "LexA protein"
XX Region 203..552
XX /note= "Human ANIC-BP-1 protein"
XX WO200170771-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 20-MAR-2001; 2001WO-EP003149.
XX
XX PR 21-MAR-2000; 2000EP-00106110.
XX
XX PA (MERE) MERCK PATENT GMBH.
XX
XX PI Den Daas I, Duecker K, Hock B;
XX
XX DR WPI; 2001-607519/69.
XX
XX PT Novel acute neuronal induced calcium binding protein type 1 ligand
XX polypeptides, useful in the treatment of stroke, head trauma, multiple
XX sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
XX injury.
XX
XX PS Disclosure; Page 44-46; 46pp; English.
XX
XX The invention relates to human acute neuronal induced calcium binding
XX protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.
XX Sequences of the invention are useful for treating human diseases
XX including stroke, head trauma, multiple sclerosis, Parkinson's disease,
XX Alzheimer's disease and spinal cord injury. They are also useful as
XX vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
XX soluble receptors. Polynucleotides of the invention are useful as
XX diagnostic reagents, for chromosome localization studies, and as valuable
XX tools for tissue expression studies. They are also useful in gene
XX therapy. The present sequence is LexA-human ANIC-BP-1 fusion protein
XX comprising the LexA protein and a C-terminally linked human ANIC-BP-1
XX protein

```
XX SQ Sequence 552 AA;
Alignment Scores:
Score: 6.52e-163 Length: 552
Percent Similarity: 94.29% Matches: 329
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AAE10859 (1-552)
QY 1 ATGCCGTTCCCGTTTGGGAAGTCTCACAATAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
DB 212 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 231
QY 61 GAGAGCATGGCTGTTCTGGAAGAACAGACATTTCTGATTAATAAGAGCAGAAAGCGCTACA 120
DB 232 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 251
QY 121 GAAGAAGTTTCCAAAATCTGTTGCCATGAAGAAATTCGTATGGCACAATAATGAATAA 180
DB 252 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 271
QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTCTGCTCAAGAACTCTATAATAGTGGCTCCTT 240
DB 272 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 291
QY 241 AGCACCTGGTAGCTGATTTACAGCTCATTTGACATTTGAGGGCAAAAAGAGCTGGCTCAA 300
DB 292 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 311
QY 301 ATTTTCAACAATATTTCTCAGAGACAAATTCGTAGCAAGTCTCTACTGTTGAATACATC 360
DB 312 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 331
QY 361 TGCACCCACAGAAATTTTGTTCATGTTATTTGAAAGGATGATGATCTCCAGAAATAGCT 420
DB 332 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 351
QY 421 CTAAATTTGGAATATTTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 352 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 371
QY 481 TTGTGCTCGGAACAGTTTTATGATTTCTTCAAGATGTCGAAATGTCACATTTGACATA 540
DB 372 LeuIrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 391
QY 541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAAGACATATAAATTCCTCAGTGCA 600
DB 392 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 411
QY 601 GAATTTTGGACACAGCATATGATAGATTTTTCAGTGAATGAGAACTTACTTCAATCA 660
DB 412 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 431
QY 661 GAAATTTATGTGACAAAAGACAGTCAGTCAAGCTTCTCGGTGAATCTACTACTAGATAGA 720
DB 432 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 451
QY 721 CACAACTTTCACAAATATGACAAATATGATCATCATGTAACCTGAGAACCTCAAAATTAATG 780
DB 452 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 471
QY 781 AACCTGCTCGGAGACAAAAGTCGCAATCCAGTTTGGAGCTTTCACGTTTTRAAGTG 840
DB 472 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 491
QY 841 TTTGTAGCAATCTTAAACAGACGAGCCCATCTCTAGACATCTCTCTCAAGAACACAGGCC 900
DB 492 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 511
```

```
901 AAACCTATAGAGTTCTCTCAGCAAGTTTTCAGAACAGACAGGACGATTTGATGACAGATTCC 960
512 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 526
961 GTACCGACGACGAATTCCTCCGGTTCGATTACGGGTTAAACCGGCGACGCGTGGGATCAGG 1020
527 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 542
1021 GATTGGAAGACAGCAGCTCAGCAAGAGCT 1050
543 AspleuLysArgProAlaGlnGlnGluAla 552

RESULT 6
AAAY94248
ID AAY94248 standard; protein; 341 AA.
XX AC AAY94248;
XX 10-AUG-2000 (first entry)
DE Mouse calcium binding protein MO25.
XX Mouse; calcium binding protein; cancer; inflammation; MO25; CBP;
XX reproductive disorder; autoimmune disorder; developmental disorder;
XX seizure disorder; immune disorder; infection.
XX Mus sp.
XX WO200029580-A1.
XX 25-MAY-2000.
XX 12-NOV-1999; 99WO-US027027.
XX 13-NOV-1998; 98US-00190965.
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;
XX WPI; 2000-387793/33.
XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
XX diagnosis, prevention and treatment of cancers, immune, developmental or
XX reproductive disorders.
XX Disclosure; Page 66-67; 72pp; English.
XX The present sequence is the mouse calcium binding protein MO25. It was
XX used in a sequence alignment to identify human calcium binding protein
XX hCBP. The hCBP protein and the gene encoding it are useful for the
XX diagnosis and treatment of the following types of disorder: cancers (such
XX as adenocarcinomas), reproductive disorders (such as infertility,
XX ovulatory defects, endometriosis, disruptions of the oestrus and
XX menstrual cycles, polycystic ovary syndrome and ovarian
XX hyperstimulation), autoimmune disorders (such as benign prostatic
XX hyperplasia and prostatitis), developmental disorders (such as Cushing's
XX syndrome, muscular dystrophy and gonadal dysgenesis), hereditary
XX neuropathies, seizure disorders, immune disorders (such as AIDS,
XX allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's
XX disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,
XX rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative
XX colitis), and viral, bacterial, fungal, parasitic, protozoal and
XX helminthic infections
XX SQ Sequence 341 AA;

Alignment Scores:
Pred. No.: 8.89e-161 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservatave: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55%
```

DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x AAY94248 (1-341)

QY 1 ATGCGGTTCCCGTTTGGGAAGCTCTCACAAATCTCCAGCAGACATTGTGGAAGAACTGTAAG 60
 |||||
 Db 1 MetProPheProPheGlySerHisLysSerProAlaAspIleValLysAsnLeuLys 20
 |||||
 QY 61 GAGAGCATGGCTGTTCTGTGAAAAACAAGACATTTCTGTATAAAAAAGCAGAAAAGCTTACA 120
 |||||
 Db 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
 |||||
 QY 121 GAAGAAGTTCCAAAATCTGGTTCGCCATGAAGAAATCTGTATGCGACAAATGAAAAA 180
 |||||
 Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
 |||||
 QY 181 GAGCCTCAGACAGAAGCAGTAGTACGTCACCTCCTCAAGAACCTCTATAATAGTGGGCTCCTT 240
 |||||
 Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
 |||||
 QY 241 AGCACCCCTGGTAGTGATTTTACAGCTCAATTGACTTTTGAGGCGCAAAAAAGACGTGGCTCAA 300
 |||||
 Db 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
 |||||
 QY 301 ATTTTCAACAATATTTCTCAGAAGACAAATTTGGTACGAGAACTCCTACTGTTGAATACATC 360
 |||||
 Db 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
 |||||
 QY 361 TGCAACCCAAACAGAAATATTTGTTTCATGTTATTGAAAGGATGATGAATCTCCAGAAATAGCT 420
 |||||
 Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
 |||||
 QY 421 CTAATTTGTGGAAATAATGTTTAAGAGAATGCATCAGACATGAACCACTTGCCAAAAATCATTT 480
 |||||
 Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
 |||||
 QY 481 TTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
 |||||
 Db 161 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
 |||||
 QY 541 GCCTTCAGATGCATTTGCCACATTTCAAGGATTTACTTACAGACATAAATTTGCTCAGTGCA 600
 |||||
 Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 |||||
 QY 601 GAATTTTGGAAACAGCATATTGATAGATTTTTTCAGTGAATATGAGAAGTTACTTTCATTCA 660
 |||||
 Db 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 |||||
 QY 661 GAAATATTATGCAAAAAGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGA 720
 |||||
 Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
 |||||
 QY 721 CACAACCTTCACAATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
 |||||
 Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 |||||
 QY 781 AACCTGCTCGAGACAAAAGTCGCAACATCCAGTTTGAAGGCTTTTCAGTTTAAAAAGTG 840
 |||||
 Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 |||||
 QY 841 TTTGTAGCCAAATCTCAACAGACGAGCCCACTCTAGACATCTCTCCTCAAGAACGAGGCC 900
 |||||
 Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
 |||||
 QY 901 AAACCTCATAGAGTTCTCAGCAAGTTTCAGAACGACGAGGACGGAATGTATGAGCAGTTCC 960
 |||||
 Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 |||||
 QY 961 GTACCGCAGCAGCAATTTCCGGGTGCAATTTACGCGTTAAACCGCGACGCGTGGATCAGG 1020
 |||||
 Db 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 |||||
 QY 1021 GATTTGAAGAGACCAGCTTCAGCAAGAAGCT 1050

Db	332 AsnLeuLysArgAlaAlaGlnGlnGluAla	341
RESULT 7		
ABG23844		
ID	ABG23844 standard; protein; 354 AA.	
XX		
AC	ABG23844;	
XX		
DT	18-FEB-2002 (first entry)	
XX		
DE	Novel human diagnostic protein #23835.	
XX		
KW	Human; chromosome mapping; gene mapping; gene food supplement; medical imaging; diagnostic	
OS	Homo sapiens.	
XX		
PN	WO200175067-A2.	
XX		
PD	11-OCT-2001.	
XX		
PF	30-MAR-2001; 2001WO-US008631.	
PR	31-MAR-2000; 2000US-00540217.	
FR	23-AUG-2000; 2000US-00649167.	
XX		
FA	(HYSE-) HYSEQ INC.	
XX		
PI	Drmanac RT, Liu C, Tang YT;	
XX		
DR	WPI; 2001-639362/73.	
DR	N-PSDB; AAS68031.	
XX		
PT	New isolated polynucleotide and encoded polynucleotide, forensics, gene mapping, and diagnostics, responsible for genetic disorders or other biodiversity.	
PT		
PT		
PT		
XX		
PS	Claim 20; SEQ ID NO 54203; 103pp; English.	
XX		
CC	The invention relates to isolated polynucleotide sequences. (I) is useful as hybridisation p reaction (PCR) primers, oligomers, and for and in recombinant production of (II). The in diagnostics as expressed sequence tags f genes. (I) is useful in gene therapy technic activity of (II) or to treat disease states useful for generating antibodies against it polypeptide in tissue, as molecular weight supplement. (II) and its binding partners a of sites expressing (II). (I) and (II) are involving aberrant protein expression or bi polypeptide and polynucleotide sequences ha diagnostics, forensics, gene mapping, ident responsible for genetic disorders or other to produce other types of data and prod amino acid sequences. ABG00010-ABG30377 rep amino acid sequences of the invention. Note patent did not appear in the printed specif electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 354 AA;	
Alignment Scores:		
Pred. No.:	5.57e-159	Length:
Scores:	1597.50	Matches:
Percent Similarity:	92.57%	Conservat
Best Local Similarity:	92.00%	Mismatche
Query Match:	88.60%	Indels:
DB:	4	Gaps:

Alignment Scores:					
Fred. No.:	5	578-159	Length:		354
Score:	1597.50	Matches:			322
Percent Similarity:	92.57%	Conservative:			2
Best Local Similarity:	92.00%	Mismatches:			17
Query Match:	88.60%	Indels:			9
DB:	4	Gaps:			2

US-10-089-688-1 (1-1053) x ABG23844 (1-354)	
QY	1 ATGCCGTTCCGTTGGAGTCTCACAATCTCCAGCAGACATTTGTAAGATCTGAAG 60
Db	14 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 33
QY	61 GAGAGCATGGCTGTTCTGGAAAGCAAGACATTTCTGATAAAAGCAGAAAGCGCTACA 120
Db	34 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 53
QY	121 GAAGAAGTTTCCAAAAATCTGGTCCCATGAAGAAATTTCTGTATGGCACAATAAGAAAA 180
Db	54 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 73
QY	181 GAGCTCAGACAGACAGTAGCTCAACTGCTCTCAAGAACTCTATATAGTGGCTCTT 240
Db	74 AspProGlnThrGluAlaGlyAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 93
QY	241 AGCACCTGGTAGCTGATTACAGCTCATTCACATTTTGAGGGCAAAAAGAGCTGGCTCAA 300
Db	94 IleThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 113
QY	301 ATTTTCAACATAATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTACTGTTGAATACATC 360
Db	114 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 133
QY	361 TGCACCCCAAGATATTTTCTCATGTTATGTAAGGGTATGATCTCCAGAAATAGCT 420
Db	134 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 153
QY	421 CTAAATTTGGAAATATCTTAAAGAGATGTCATCAGACATGAACCACTTGCAGAAATCATT 480
Db	154 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuGlyLysIle 173
QY	481 TTGGTTCGGAACAGTTTTATGATTTCTCAGATATGTCGAAATGTCACAAATTTGACATA 540
Db	174 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 193
QY	541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACACACATTAATTTCTCAGTGCA 600
Db	194 AlaSerAspAlaPheAlaThrPheLysGlyLeuLeuThrArgHisLysLeuLeuSerAla 213
QY	601 GAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAGCTTACTTCAATCA 660
Db	214 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 233
QY	661 GAAATATATGACAAAAAGACAGTCACTAGCTTCTCGGTGAACCTACTACTAGATAGA 720
Db	234 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 253
QY	721 CACAACTTCACATTTATCACAATAATACATCAGTAACTGAGAACTTCAAAATTTAATGATG 780
Db	254 HisAsnPheThrIleMetThrLysTyrIleSerLysProValAsnLeuLysLeuMetMet 273
QY	781 AACCTGCTGCAGACAAAGTCGCAACATTCAGATTTGAGGCTTTCACGTTTTTAAAGTG 840
Db	274 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysAla 293
QY	841 TTTGTAGCCAATCTTCAACAGACGAGCCCATCTCTAGACATCTCTCTCAAGAACGAGCC 900
Db	294 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 313
QY	901 AAACATAGAGTTCTCTCAGCAAGTTTTCAGACGACAGCGAGGATTTGTATGACGATCC 960
Db	314 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 328
QY	961 GTACCGCAGCAATCTCCGGGTTCGATTTACCGGTTTAAACCGCGACGCTGGATCAGG 1020
Db	329 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 344
QY	1021 GATTTCAAGACAGACAGCTCAGCAAGAGCT 1050
Db	345 AspLeuLysArgProAlaGlnGlnGluAla 354

RESULT 8

AA94247 standard; protein; 337 AA.

AA94247;

10-AUG-2000 (first entry)

Human calcium binding protein hCBP.

Human; calcium binding protein; cancer; inflammation; CBP; reproductive disorder; autoimmune disorder; developmental disorder; seizure disorder; immune disorder; infection.

Homo sapiens.

WC200029580-A1.

25-MAY-2000.

12-NOV-1999; 99WO-US027027.

13-NOV-1998; 98US-00190965.

(INCY-) INCYTE PHARM INC.

Tang YT, Guegler KJ, Corley NC, Gorgone GA;

WPI; 2000-387793/33.

N-PSDB; AAA27332.

Human hCBP protein, and the nucleic acid encoding it, useful for e.g. diagnosis, prevention and treatment of cancers, immune, developmental or reproductive disorders.

Claim 1; Fig 1; 72pp; English.

The present sequence is the human calcium binding protein hCBP. It was obtained by screening a coronary artery smooth muscle cDNA library, from which five overlapping nucleic acids were isolated, sequenced and expressed to give the protein. The protein and the gene encoding it are useful for the diagnosis and treatment of the following types of disorder: cancers (such as adenocarcinomas), reproductive disorders (such as infertility, ovulatory defects, endometriosis, disruptions of the oestrus and menstrual cycles, polycystic ovary syndrome and ovarian hyperstimulation), autoimmune disorders (such as benign prostatic hyperplasia and prostatitis), developmental disorders (such as Cushing's syndrome, muscular dystrophy and gonadal dysgenesis), hereditary neuropathies, seizure disorders, immune disorders (such as AIDS, allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves' disease, multiple sclerosis, psoriasis, rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative colitis), and viral, bacterial, fungal, parasitic, protozoal and helminthic infections

Sequence 337 AA;

Alignment Scores:	2.28e-127	Length:	337
Pred. No.:	1297.50	Matches:	263
Score:	85.26%	Conservative:	32
Percent Similarity:	76.01%	Mismatches:	38
Best Local Similarity:	71.96%	Indels:	13
Query Match:	3	Gaps:	3

US-10-089-688-1 (1-1053) x AA94247 (1-337)

QY 1 ATGCCGTTCCGTTGGAGTCTCACAATCTCCAGCAGACATTTGTAAGATCTGAAG 60

Db 4 MetProLeu---PheSerLysSerHisLysSerProAlaGluIleValLysLeuLys 22

QY 61 GAGAGCATGGCTGTTCTGGAAAGCAAGACATTTCTGTATATAAAGCAGAAAGCGCTACA 120

```

Db      23  AsnLeuAlaLeuGluGln-----AspLysLysThrAspLysAlaSer 39
Qy      121 GAAGAAGTTTCCAAATCTGGTCCCATGAAAGAAATCTGTATGGCACAATGAAAAA 180
Db      40  GluGluValSerLysSerLeuGlnAlaMetLysGluLeuLysGlyThrAsnGluLys 59
Qy      181 GAGCTCAGACAGACAGTAGCTCACTGCTCAAGAACTCTATATAGTGGCTCCTT 240
Db      60  GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrSerSerGlyLeuLeu 79
Qy      241 AGCACCTGGTAGCTGATTACAGCTCATTGACCTTTGAGGGCAAAAAGAGCTGGCTCAA 300
Db      80  ValThrLeuLeuAlaAspLeuGlnLeuLeuAspPheGluGlyLysLysAspValThrGln 99
Qy      301 ATTTTCAACATATCTCAGAAAGCAAAATGGTACGAGAACTCCTACTGTTTGAATACATC 360
Db      100 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
Qy      361 TGCACCCCAACAGATATTTGTTCTGATGTTTATGAGGGTATGATCTCCAGAAATAGCT 420
Db      120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
Qy      421 CTAATTTGTGAATAATGTTAAGAGATGTCATCAGACATGAACCACTTGCACAAATCATT 480
Db      140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
Qy      481 TTGTGGTCGGAACAGTTTATGATTCTTCATGATATGTCGAAATGTCAACATTTGACATA 540
Db      160 LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
Qy      541 GCTTCAGATGCTATGTCACAGATTCAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
Db      180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
Qy      601 GAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAGTACTTCAATCA 660
Db      200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
Qy      661 GAAATATATGTGACAAAGACAGTCACTGAAGCTTCTCGGTGAACACTACTAGATAGA 720
Db      220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 239
Qy      721 CACAATCTCAATATGACAAATATACATCAGTAAACCTGAGAACCTCAATTAATGATG 780
Db      240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
Qy      781 AACCTGCTGCAGACAAAGTCGCAACATCCAGTTTGAGGCTTTTCAGCTTTTAAAGTG 840
Db      260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
Qy      841 TTTGTAGCCAACTCTAAACAGACGACGCCCATCTCAGACATCCCTCCTCAAGAACCCAGGCC 900
Db      280 PheValAlaSerProHisLysThrGlnProIleValGluLeuLeuLysAsnGlnPro 299
Qy      901 AAATCTATAGATGCTCCTCAGCAAGTTTCAGAACGACGAGGAGATTTGATGAGCAGTTCC 960
Db      300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
Qy      961 GTACCGACGAGATTCCTGGGTGATTACGCGTTAAACCGCGGACGCGTGGGATCAGG 1020
Db      315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
Qy      1021 GATTTGAAGACACAGCT 1038
Db      331 AspLeuLysLysThrAla 336

RESULT 9
AAB82090
ID AAB82090 standard; protein; 337 AA.
XX AC
AC AAB82090;
XX
```

```

DT      26-JUN-2001 (first entry)
DE      Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP.
KW      Human; cerebroprotective; neuroprotective; vulnerable; vaccine;
KW      Gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW      stroke; acute head trauma; multiple sclerosis; spinal cord injury.
XX      Homo sapiens.
OS      WO200123552-A1.
PN      05-APR-2001.
XX      18-SEP-2000; 2000WO-EP009132.
XX      24-SEP-1999; 99EP-00118848.
PA      (MERE ) MERCK PATENT GMBH.
PI      Den Daas I, Duecker K;
XX      WPI; 2001-308142/32.
DR      N-PSDB; AAF86462.
XX      Novel human acute neuronal induced calcium binding polypeptide, and
PT      polynucleotides encoding them useful for diagnosing or treating stroke,
PT      acute head trauma, multiple sclerosis and spinal cord injury.
XX      Claim 1; Page 41-42; 45pp; English.
XX      The present sequence is the protein sequence for human Acute Neuronal
CC      Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
CC      protein are useful for treating stroke, acute head trauma, multiple
CC      sclerosis and spinal cord injury. ANIC-BP coding sequence and protein are
CC      also useful as vaccines for inducing an immunological response in a
CC      mammal
XX      SQ Sequence 337 AA;

Alignment Scores:
Pred. No.: 2.28e-127 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: Gaps: 3

US-10-089-688-1 (1-1053) x AAB82090 (1-337)
Qy      1 ATCCGTTCCCGTTTGGGAAGTCTCAAAATCTCCAGACAGCATTTGTGAAGAATCTGAAG 60
Db      4 MetProLeu--PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
Qy      61 GAGAGCATGCTGTTCTGGAAGAACAGACATTTCTGATATAAAGACAGAAAGGCTACA 120
Db      23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
Qy      121 GAAGAAGTTTCCAAATCTGGTTCATCAAGAAATTTCTGTATGGCACAATGAAAAA 180
Db      40 GluGluValSerLysSerLeuGlnAlaMetLysGluLeuLysGlyThrAsnGluLys 59
Qy      181 GAGCTCAGACAGACAGTAGCTCACTGCTCAAGAACTCTATATAGTGGCTCCTT 240
Db      60 GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrSerSerGlyLeuLeu 79
Qy      241 AGCACCTGGTAGCTGATTACAGCTCATTGACCTTTGAGGGCAAAAAGAGCTGGCTCAA 300
Db      80 ValThrLeuIleAlaAspLeuGlnLeuLeuAspPheGluGlyLysLysAspValThrGln 99
Qy      301 ATTTTCAACATATTTCTCAGAAAGCAAAATGGTACGAGAACTCCTACTGTTTGAATACATC 360
Db      100 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
```

```
QY 361 TGCACCCACAGAAATATTGTTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
QY 421 CTAATTTGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAGAAATCATT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
QY 481 TTGTGTCGGAACAGTTTATGATTCTTTCAGATATGTCGAAATGTCACATTTGCACATA 540
Db 160 LeuPheSerAnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
QY 541 GCTTCAGATGATTGTCACATTCAGGATTACTTACAGACACATAAATCTCAGTGCA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY 601 GAATTTTGGACAGCATATATGATAGATTTTTCAGTGAATATGAGAAATGTTACTTCATCA 660
Db 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuGlnSer 219
QY 661 GAAATATATGTCACAAAAGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGTAGA 720
Db 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAACTTCACAATATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
Db 240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACCTGCTGCAGACAAAAGTCGCAACATCCAGTTTTCAGGCTTTTCAGCTTTTAAAGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCAATCTTAACAGACGAGCCCATCTAGACATCTCTCTCAAGAACCCAGGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
QY 901 AAATCATATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGACGAGTGTGATGAGCAGTCC 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY 961 GTACCGACGACGAATCTCCGGTTCGATTTCACGCGTTAAACCGGACGCGTGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
QY 1021 GATTGAAGACACAGCT 1038
Db 331 AspLeuLysLysThrAla 336

RESULT 10
AAM39078
ID AAM39078 standard; protein; 337 AA.
XX
AC AAM39078;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2223.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
```

```
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-0055317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI58234.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
PS Example 4; SEQ ID NO 2223; 10078pp; English.
XX
The invention relates to human nucleic acids (AAI57798-AAI61369) and the
encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders. Note: The sequence data for this patent did not form
part of the printed specification
XX
SQ Sequence 337 AA;
Alignment Scores:
Pred. No.: 2,28e-127 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: Gaps: 3
US-10-089-688-1 (1-1053) x AAM39078 (1-337)
QY 1 ATGCCGTTCCCGTTTGGGAAGTCTCACAAATCTCCAGCAGCATTTGGAAGATCTGAG 60
Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATCTTCTGATAAAGCAAGCAAGAAAGCTACA 120
Db 23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
QY 121 GAAGAAGTTTCCAAAAATCTGTTGCCATGAAGAAATCTGTATGGCACAATGAAAAA 180
Db 40 GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
QY 181 GAGCCTCAGACAGAGCAGTAGCTCAACTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuLysTyrSerSerGlyLeuLeu 79
QY 241 AGCACCTGGTAGCTGATTACAGCTCATTGAGGCGCAAAAGAGCTGGCTCAA 300
Db 80 ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
```

```
QY 301 ATTTCAACAATATCTCAGAGCAAAATTTGGTACGAGAACTCTCTACTGTTGATACATC 360
Db 100 ILePheAsnAsnIleLeuArg-gglnleGlyThrArgSerProThrValGluTyrIle 119
QY 361 TGCACCCAAACAGAAATATTTGTTTCATCTTATTTGAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
QY 421 CTAAATTTGGAATATATTTAAAGAAATGTCATCAGACATGAACACATCTGCAAAATCATTT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
QY 481 TTGTGGTCGGAACAGTTTATTTCTTCAGATATGTCGAAATGTCAAACATTTGCACATA 540
Db 160 LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
QY 541 GCTTCAGATGCATTTGTCACATTCAGGATTTACTTTACAGACATAAATGCTCAGTGCA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY 601 GAATTTTGGAAAGCATTTATGATAGATTTTTCAGTGAATATGAGAGTACTTCAATTC 660
Db 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
QY 661 GAAATTTATGTCACAAACACAGTCACTCAAGCTTCTCGTCAACTACTACTAGATAGA 720
Db 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAACCTTCACAATATATGACAAATATACATCAGTAAACCTCAGAACCTCAAAATTAATGATG 780
Db 240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACTGCTGGAGACAAAGTCGCAATCCAGTTTGGAGCCTTTTACGTTTAAAGGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCCCAATCTTAACAGAGCGCCCATCTAGACATCTCTCTCAAGACCGGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
QY 901 AAACATCATAGAGTCTCTCAGCAAGTTTTCAGAACGACAGGACGATTTGTATGACAGTTC 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY 961 GTACCGACGACGAATCCCGGGTCGATTTACGCGTTTAAACCGCGACGCGTGGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
QY 1021 GATTTGAGAGACGACGCT 1038
Db 331 AspLeuLysLysThrAla 336

RESULT 11
AAB94139
ID AAB94139 standard; protein; 289 AA.
XX
AC AAB94139;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:14408.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
```

```
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 14408; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a polynucleotide which comprises a 5'-end
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and
CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAB13629 to AAB13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 289 AA;

Alignment Scores:
Pred. No.: 3,23e-112 Length: 289
Score: 1153.50 Matches: 231
Percent Similarity: 85.52% Conservative: 23
Best Local Similarity: 77.78% Mismatches: 34
Query Match: 63.98% Indels: 9
DB: 4 Gaps: 1

US-10-089-688-1 (1-1053) x AAB94139 (1-289)
QY 148 ATGAAGAATAATCTGTATGGCAAAATGAAAGAGCCCTCAGACAGACAGTACTCA 207
Db 1 MetLysGluIleLeuCysGlyThrAsnGluLysGluProThrGluAlaValAlaGln 20
QY 208 CTTGCTCAAGAACTCTATAATAGTGGCTCCCTTAGCACCCCTGGTACCTGATTTACAGCTC 267
Db 21 LeuAlaGlnGluLeuTyrSerGlyLeuLeuValThrLeuIleAlaAspLeuGlnLeu 40
QY 268 ATTGACTTTGAGCGCAAAAGACGCTGGCTCAAAATTTTCAACAATATTTCTCAGAGACAA 327
Db 41 IleAspPheGluGlyLysLysAspValThrGlnIlePheAsnAsnIleLeuArgGln 60
QY 328 ATTGTACAGAACTCTACTACTCTTGAATACATCTCGACCCCAACAGAAATATTTGTTTCATG 387
Db 61 IleGlyThrArgSerProThrValGluTyrIleSerAlaHisProHisIleLeuPheMet 80
QY 388 TTATTGAAGGCTATGATCTCCAGAAATAGCTTAAATTTGGATATATGTTAAGAA 447
Db 81 LeuLeuLysGlyTyrGluAlaProGlnIleAlaLeuArgCysGlyIleMetLeuArgGlu 100
```

Qy	448	TGATCAGACATGACCACTTGCAGAAATCATTTTGTGGTCGGAACAGTTTATGATTC	507
Db	101	Cysilearghisgluproleuvallyleilleleupheserasnleupheargasp	120
Qy	508	TTGATATATGCAATGTCAACATTTGACATAGCTTCAGATGCATTTGCCACATTCAG	567
Db	121	Phelyslyrvalgluleuserthrphesaspilalaseraspalahealathrphelys	140
Qy	568	GATTACTTCAAGACATAAATTCCTCAGTCGCAAAATTTTGGACAGCAATATGATAGA	627
Db	141	AspleuLeuthrarghislysalLeuValalaaspheleuGluGlnAsenTyAspThr	160
Qy	628	TTTTTTCAGTGAATGAGAGTTACTTCATTCAGAAATTTATGTCAGAAAACAGATCA	687
Db	161	IlePheGluAspTyrgluysleuLeuGlnSerGluAsnTyrgluValThrlysrargGlnSer	180
Qy	688	CTGAAGCTTCTCGTGAACTACTACTAGATAGACACAACTTCAATTTATGACAAAATAC	747
Db	181	LeuylsleuLeuGlyGluLeuileLeuasparghisasnPhealalleMetThrlysrTy	200
Qy	748	ATCAGTAACTGAGAACCTCAATTAATGATCAACCTGTCGAGACAAAAGTCGCAAC	807
Db	201	IleserlysrProGluAsnLeuylsLeuMetMetAsnLeuLeuargAspLysSerProasn	220
Qy	808	ATCAGTTTGGGCTTTCAGTTTAAAGTGTGTGAGCCAACTCCTAACAGACGACG	867
Db	221	IleglnPheGluAlaPheHisValPheLyValPheValalaSerProHislysrThrgln	240
Qy	868	CCCATCTGACATCTCTCAAGAACAGCCCAAACTCATAGAGTTCTCTCAGCAAGTTT	927
Db	241	ProlevalGluilleLeuLeuylsasnGlnProlyslleulleGluPheLeuSerSerPhe	260
Qy	928	CAGAACAGCAGGACGATTTGATGACGATTCGTCACGACGAAATTCGCGGTGCAT	987
Db	261	GlnlysrGluArgThrAsp-----AspGluGlnPheAla	271
Qy	988	TTACGCGTTAAACCGCGGCGTGGGATCAGGATTTGAGACACCGCT	1038
Db	272	AspGluylsAsenTyrgluilleLysGlnleargaspLeuylslysrAla	288
RESULT 12			
ID	ABB60392	standard; protein; 339 AA.	
XX	AC		
XX	ABB60392;		
DT	26-MAR-2002	(first entry)	
DE	Drosophila melanogaster	polypeptide SEQ ID NO 7968.	
KW	Drosophila;	developmental biology; cell signalling; insecticide;	
KW	pharmaceutical.		
OS	Drosophila melanogaster.		
FN	WO200171042-A2.		
PD	27-SEP-2001.		
XX	23-MAR-2001;	2001WO-US009231.	
XX	23-MAR-2000;	2000US-0191637P.	
PR	11-JUL-2000;	2000US-00614150.	
XX	(PEKE)	PE CORP NY.	
XX	Venter JC,	Adams M, Li PWD, Myers EW;	
XX	WPI;	2001-656860/75.	
DR	N-PSDB;	ABL04495.	
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more	

genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

Disclosure; SEQ ID NO 7968; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16178-ABL30511), expressed DNA sequences (AB101840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 339 AA;

Alignment Scores:

aligned. No.:	5.63e-109	Length:	339
core:	1123.00	Matches:	221
Percent Similarity:	84.95%	Conservative:	50
Best Local Similarity:	69.28%	Mismatches:	40
Query Match:	62.29%	Indels:	8
DB:	4	Gaps:	5

US-10-089-688-1 (1-1053) x ABB60392 (1-339)

Qy	1	ATGCCGTTCCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGAGAAATCTGAAG	60
Db	1	MetProLeu---PheGlyLysSerGlnLysSerProValGluLeuValLysSerLeuLys	19
Qy	61	GAGCAGATGCGTCTTCTGGAAAAGCAAGACATTTCTGTATAAAAAAGCAGAAAAGGCTACA	120
Db	20	GluAlaIleasnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln	36
Qy	121	GAAGAAGTTTCCAAAATCTGTTGCCATGAAAGAAATCTGTATGGCACAATAAGAAAA	180
Db	37	GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuTyrglySerSerAspAla	56
Qy	181	GAGCCTCAGACAGAA---GCAGTAGTCTCACTTGCTCGAAGACTCTATAATAGTGGGCTC	237
Db	57	GluProProAlaAspTyrgluValAlaGlnLeuSerGlnGluLeuTyrglnSerAsnLeu	76
Qy	238	CTTAGCACCTGCTGAGTGTATTTACAGCTCATTGACCTTTGAGGCGCAAAAAGCGTGGCT	297
Db	77	LeuLeuLeuLeuIleGlnasnLeuHisArgIleaspPheGluGlyLysLysHisValala	96
Qy	298	CAAAATTTCAACAATATTTCTCAGACAGACAAATTTGGTACGAGAACTCTCTACTGTTGAATAC	357
Db	97	LeuilePheAsnAsnValLeuArgGlnleGlyThrArgSerProThrValGluTyrg	116
Qy	358	ATCTGACCCACAGAAATATTTTGTTCATGTTTATTGAAAGGGTATGAACTCT-----CCA	411
Db	117	IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlyTyrgluAspAlaHisPro	136
Qy	412	GAAATAGCTCTAAATTTGTGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCA	471
Db	137	GluileAlaLeuasnSerGlyThrMetLeuargGluCysalaargTyrgluAlaLeuAla	156
Qy	472	AAATCAATTTTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACA	531
Db	157	LysileMetLeuHisSerAspGluPhePheLysPhePheArgTyrgluValGluValSerThr	176
Qy	532	TTTGACATAGCTTCAGATGCAATTTGCCACATTTGCCACATTTTACTTACAAGACATAAATG	591
Db	177	PheAspIleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeu	196
Qy	592	CTCAGTCAGAAATTTTGGAAACAGCATTTATCATGATTTTTCAGTGAA---TATGAGAAG	648
Db	197	LeuCysalaGluPheLeuAspAlaAsnTyrglnAspLysPhePheSerGlnHisTyrglnArg	216
Qy	649	TTACTTCATTCAGAAAATTTATGTGACAAAAGACAGCTCAGTGAAGCTTCTCGGTGAAC	708

Db 297 ArgAsnGlnThrLysLeuValAspPheLeuThrAsnPheHisThrAspArgSerGlu 315

RESULT 14

AA94250

ID AAY94250 standard; protein; 377 AA.

XX

AC AAY94250;

XX

DT 10-AUG-2000 (first entry)

XX

DE C. elegans yeast-like calcium binding protein.

XX

KW Calcium binding protein; cancer; inflammation; yeast-like CBP; CBP;

KW reproductive disorder; autoimmune disorder; developmental disorder;

KW seizure disorder; immune disorder; infection.

XX

OS Caenorhabditis elegans.

XX

PN WO200029580-A1.

XX

PD 25-MAY-2000.

XX

PF 12-NOV-1999; 99WO-US027027.

XX

PR 13-NOV-1998; 98US-00190965.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX

DR WPI; 2000-387793/33.

XX

PT Human hCBP protein, and the nucleic acid encoding it, useful for e.g.

PT diagnosis, prevention and treatment of cancers, immune, developmental or

PT reproductive disorders.

XX

PS Disclosure; Page 68-69; 72pp; English.

XX

CC The present sequence is the C. elegans yeast-like CBP. It was used in a

CC sequence alignment to identify human calcium binding protein hCBP. The

CC hCBP protein and the gene encoding it are useful for the diagnosis and

CC treatment of the following types of disorder: cancers (such as

CC adenocarcinomas), reproductive disorders (such as infertility, ovulatory

CC defects, endometriosis, disruptions of the oestrus and menstrual cycles,

CC polycystic ovary syndrome and ovarian hyperstimulation), autoimmune

CC disorders (such as benign prostatic hyperplasia and prostatitis),

CC developmental disorders (such as Cushing's syndrome, muscular dystrophy

CC and gonadal dysgenesis), hereditary neuropathies, seizure disorders,

CC immune disorders (such as AIDS, allergies, anaemia, asthma,

CC atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'

CC disease, multiple sclerosis, psoriasis, rheumatoid arthritis,

CC scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,

CC bacterial, fungal, parasitic, protozoal and helminthic infections

XX

SQ Sequence 377 AA;

Alignment Scores:

Pred. No.:	9,71e-102	Length:	377
Score:	1054.50	Matches:	205
Percent Similarity:	77.44%	Conservative:	49
Best Local Similarity:	62.50%	Mismatches:	61
Query Match:	58.49%	Indels:	13
DB:	3	Gaps:	2

US-10-089-688-1 (1-1053) x AAY94250 (1-377)

QY 1 ATGCCGTTCCCGTTGGGAAGTCTCACAATCTCCAGCAGACATTGTGAAGAACTGAAG 60

Db 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20

QY 61 GAGACATGGCTGTCTCGAAAG-----CAAGAC 90

Db 21 GluValLeuThrIleLeuAspLysLeuProProProLysLeuAspLysAspGlyAsnIle 40

QY 91 ATTTCTGATAAAAAGCAGAAAAGGCTACAGAAGAGTTTCCAAAATCTGGTGGCCATG 150

Db 41 GlnSerAspLysLysTyAspLysAlaLeuAspLysValSerLysAsnValAlaMetIle 60

QY 151 AAAGAAATTTCTGTATGCGACAAATGAAAAGAGCCTCAGACAGAA-----GCAGTA 201

Db 61 LysSerPheIleTyGlyAsnAspSerAlaGluProSerSerGluHisValValGlnVal 80

QY 202 GCTCAACTTCTCAGAACTCTATATAGTGGGCTCTTAGCACCCTCGGTAGCTGATTTA 261

Db 81 AlaGlnLeuAlaGlnGluValTyAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100

QY 262 CAGCTCATCTTGTGAGGCGCAAAAGACGTGGCTCAAATTTTCAACAATATTCTCAGA 321

Db 101 ProLysPheGluPheGluCysLysAspValGlyGlnIlePheAsnAsnLeuLeuArg 120

QY 322 AGACAAATTTGTTAGAGAACTCTTACTGTTGAATACATCTGCACCCCAAGAAATATTTG 381

Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyLeuGlyAlaArgProGluIleLeu 140

QY 382 TTCATGTTATTGAAGGTTATGAATCTCCAGAATAGCTCTAAATTTGTGGAATATCTTA 441

Db 141 IleGlnLeuValGlnGlyTySerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160

QY 442 AGAAGATGCATCAGACATGAACCACTTGCAAAATCATTTTGTGGTGGGAACAGTTTAT 501

Db 161 ArgGluSerIleArgHisAspHisLeuAlaIleIleLeuTySerAspValPheTy 180

QY 502 GATTTCTTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGATTTGCCA 561

Db 181 ThrPheLeuTyValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200

QY 562 TTCAGAGATTTACTTCAACAGACATTAATTCCTCAGTGCAGAAATTTTGGACAGCATTA 621

Db 201 PheLysGluLeuThrThrArgHisLysAlaIleIleAlaGluPheLeuAspSerAsnTy 220

QY 622 GATAGATTTTTCAGTGAATATGAGAGTTACTTTCATTTCAGAAAATTTATGTCAAAAAGA 681

Db 221 AspThrPheAlaGlnTyGlnAsnLeuLeuAsnSerLysAsnTyValThrArgArg 240

QY 682 CAGTCACTGAAGTTCTCGGTGAACCTACTACTAGATAGACACAACTTCACAAATATGACA 741

Db 241 GlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArgHisAsnPheAsnThrMetThr 260

QY 742 AATATACATCATGAACCTGAGAACCTCAATTAATGATGAACCTGCTGCAGACAAAGT 801

Db 261 LysTyrlSerAsnProAspAsnLeuArgLeuMetGluLeuLeuArgAspLysSer 280

QY 802 CGCAACATCCAGTTTGAAGGCTTTTCAGTGTTTTAAAGTGTTTGTAGCCAAATCCTAACAG 861

Db 281 ArgAsnIleGlnTyGluAlaPheHisValPheLysValPheValAlaAsnProAsnLys 300

QY 862 ACGAGGCCATCTCAGACATCTCCCTCAAGAACCGGCCAACTCATAGATTCCTCAGC 921

Db 301 ProLysProIleSerAspIleLeuAsnArgAsnArgLysLeuValGluPheLeuSer 320

QY 922 AAGTTTCAGAACGACGACGACGAT 945

Db 321 GluPheHisAsnAspArgThrAsp 328

RESULT 15

AA40864

ID AA40864 standard; protein; 237 AA.

XX

AC AA40864;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 5795.

XX

KW Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

Homo sapiens.

W0200153312-Al.

26-JUL-2001.

26-DEC-2000; 2000WO-US034263.

23-DEC-1999; 99US-00471275.

21-JAN-2000; 2000US-00488725.

25-APR-2000; 2000US-00552317.

20-JUN-2000; 2000US-00598042.

19-JUL-2000; 2000US-00620312.

03-AUG-2000; 2000US-00653450.

14-SEP-2000; 2000US-00662191.

19-OCT-2000; 2000US-00693036.

29-NOV-2000; 2000US-00727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Aseundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Pi J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;

WPI; 2001-442253/47.

N-PSDB; AAI60020.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Example 2; SEQ ID NO 5795; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

Sequence 237 AA;

Alignment Scores:
Pred. No.: 2.55e-84 Length: 237
Score: 888.50 Matches: 178
Percent Similarity: 83.54% Conservative: 20
Best Local Similarity: 75.11% Mismatches: 34
Query Match: 49.28% Indels: 5
DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AAM40864 (1-237)

334 ACAGAACTCTACTGTGTAATACATCTGCACCAACAGAAATATTTGTCATGTTATTG 393

2 ThrArgSerProThrValGluTyrIleSerAlaHisProHisIleLeuPheMetLeuLeu 21

394 AAGGGTATGATCTCCAGAAATAGCTCTAAATGTCGAATATGTTAAGAGATGCATC 453

22 LysGlyTyrGluAlaProGlnIleAlaLeuArgCysGlyIleMetLeuArgGluCysIle 41

QY 454 AGACATGAACCACTTGCGAAAAATCATTTTGGTGGAGACAGTTTATGATTTCTTTCAGA 513
DB 42 ArgHisGluProLeuAlaLysIleIleLeuPheSerAsnGlnPheArgAspPhePheLys 61
QY 514 TATGTCGAAATGTCAACATTTGACATAGCTTTCAGATGATTTGGCACATTCACAGGATTTA 573
DB 62 TyrValGluLeuSerThrPheAspIleAlaSerAspAlaPheAlaThrPheLysAspLeu 81
QY 574 CTTACAGACATAAATTTGCTCAGTGCAGATTTTGGACAGCAGCATTTATGATAGATTTTTC 633
DB 82 LeuThrArgHisLysValLeuValAlaAspPheLeuGluGlnAsnTyrAspThrIlePhe 101
QY 634 AGTGAATATGAGAAGTTTACTTTCATTTCAGAAAAATTTATGTCAGAAAAACAGCAGTCTAGAG 693
DB 102 GluAspTyrGluLysLeuLeuGlnSerGluAsnTyrValThrLysArgGlnSerLeuLys 121
QY 694 CTTCTCGGTGAACCTACTACTAGATAGACACAACTTTCACAATTATGACAAAAATACATCAGT 753
DB 122 LeuLeuGlyGluLeuIleLeuAspArgHisAsnPheAlaIleMetThrLysTyrIleSer 141
QY 754 AAACCTGAGAACCTCAAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACATCCAG 813
DB 142 LysProGluAsnLeuLysLeuMetAsnLeuLeuArgAspLysSerProAsnIleGln 161
QY 814 TTTGAGGCTTTTCAAGTTTAAAGTGTGTTAGTCCAAATCCTAACAGACGACGCCCATC 873
DB 162 PheGluAlaPheHisValPheLysValPheValAlaSerProHisLysThrGlnProIle 181
QY 874 CTAGACATCTCTCAAGACACCGCCCAACTCATAGAGTTCTCTCAGCAAGTTTCAGAAC 933
DB 182 ValGluIleLeuLeuLysAsnGlnProLysLeuIleGluPheLeuSerSerPheGlnLys 201
QY 934 GACAGACGAGTGTATGAGCAGTTTCGTACCGACGACGAATTCCTCCGGGTC----- 984
DB 202 GluArgThrAsp---AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIle 220
QY 985 ---GATTACGCGTTAAACCGCGGACGCGTGGGATCAGGATTTGAAGAGA 1032
DB 221 ArgAspLeuLysLysThrAlaPro***ArgAlaLeuArgAspSerLysArg 237

Search completed: November 10, 2004, 20:43:03
Job time : 180.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:34:48 ; Search time 36.5 Seconds
(without alignments)
3826.458 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atgcgcgtccgtttggaa.....cagctcgaagaagcttaa 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Deiop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model
-Q=/cgn2_1/USPTO_pool_P/US10089688/runat_10112004_152247_4873/app_query.fasta_1.1223
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089688 @CGN_1_1_46 @runat_10112004_152247_4873 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/PCFUS COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1614.5	89.5	341	3	US-09-190-965-3
2	1614.5	89.5	341	3	US-09-470-253-3
3	1297.5	72.0	337	3	US-09-190-965-1
4	1297.5	72.0	337	3	US-09-470-253-1
5	1117	62.0	339	3	US-09-190-965-4
6	1117	62.0	339	3	US-09-470-253-4
7	1054.5	58.5	377	3	US-09-190-965-5
8	1054.5	58.5	377	3	US-09-470-253-5
9	358.5	19.9	165	4	US-09-248-796A-14303
10	166	9.2	125	4	US-09-248-796A-15202
11	163	9.0	90	4	US-09-248-796A-15201
12	115.5	6.4	630	4	US-09-248-796A-20275

13	113	6.3	541	4	US-09-134-000C-5420	Sequence 5420, Ap
14	113	6.3	2184	4	US-09-417-485D-6	Sequence 6, Appli
15	110.5	6.1	533	3	US-09-134-001C-4053	Sequence 4053, Ap
16	106	5.9	3959	3	US-08-970-269A-30	Sequence 30, Appl
17	106	5.9	3959	3	US-09-407-562-30	Sequence 30, Appl
18	105.5	5.9	3878	4	US-09-914-259-11	Sequence 11, Appl
19	105	5.8	478	4	US-09-328-352-6512	Sequence 6512, Ap
20	105	5.8	496	4	US-09-543-681A-6485	Sequence 6485, Ap
21	105	5.8	627	1	US-09-248-796A-20614	Sequence 20614, A
22	105	5.8	776	1	US-08-021-601-2	Sequence 2, Appli
23	105	5.8	776	1	PCT-US94-01624-2	Sequence 2, Appli
24	105	5.8	776	5	US-09-248-796A-20513	Sequence 20513, A
25	104	5.8	933	3	US-09-572-191-2	Sequence 2, Appli
26	103.5	5.7	1388	3	US-09-723-262-2	Sequence 2, Appli
27	103.5	5.7	1388	3	US-09-723-219-2	Sequence 3, Appli
28	103.5	5.7	2474	4	US-08-265-967C-3	Sequence 3, Appli
29	103.5	5.7	2474	4	US-08-305-790B-4	Sequence 4, Appli
30	103.5	5.7	2474	4	US-09-543-681A-6286	Sequence 6286, Ap
31	103	5.7	1155	4	US-08-685-576-4	Sequence 4, Appli
32	103	5.7	1388	2	US-09-134-001C-2940	Sequence 2940, Ap
33	102.5	5.7	670	3	US-09-107-532A-5994	Sequence 5994, Ap
34	102	5.7	506	4	US-09-543-681A-4262	Sequence 4262, A
35	101.5	5.6	467	4	US-09-248-796A-19030	Sequence 19030, A
36	101	5.6	489	4	US-08-685-576-1	Sequence 1, Appli
37	100.5	5.6	1388	2	US-08-942-012B-33	Sequence 33, Appli
38	100	5.5	460	3	US-09-269-861A-8	Sequence 8, Appli
39	100	5.5	831	4	US-09-150-867-1	Sequence 1, Appli
40	99	5.5	2954	4	US-09-710-279-2482	Sequence 2482, Ap
41	98.5	5.5	568	4	US-09-543-681A-6407	Sequence 6407, Ap
42	98.5	5.5	967	4	US-09-710-279-1432	Sequence 1432, Ap
43	98.5	5.5	1145	4	US-09-134-001C-3242	Sequence 3242, Ap
44	98.5	5.5	1151	3	US-09-976-594-296	Sequence 296, App
45	98	5.4	1388	4		

ALIGNMENTS

RESULT 1
US-09-190-965-3
; Sequence 3, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PP-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRI
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-190-965-3

Alignment Scores:
Pred. No.: 8.34e-173 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-190-965-3 (1-341)

QY 1 ATGCGTTCCTCCGTTGGAGTCTCACAAATCTCCACAGACATTGTGAGAACTCTGAAG 60
|||||
Db 1 MetcProPheProGlyLysSerHisLysSerProAlaSerPheValLysAsnLeuLys 20


```
Db 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer 220
Qy 661 GAAATATTATGTGACAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
Qy 721 CACAACTTCACAATTATGACAAAATACATCATGATAAAGCTGAGAACCTCAAAATTAATGATG 780
Db 241 HisAsnPhethrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
Qy 781 AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTTCAGGCGCTTTCACGTTTTTAAAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTTGTAGCCATCTTACACAGACGAGCCCATCTCTAGACATCTCTCCCTCAAGAACACGAGCC 900
Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAACCTCATAGAGTTCCCTCAGCAAGCTTTCAGAACACGAGGAGCGGATTGTATGAGCAGTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu-----IleArg 331
Qy 961 GTACCGACGAGCAATTCCCGGTTCGATTTCACGCTTAAACCGCGAGCGGTGGATCAGG 1020
Db 331 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331

RESULT 3
US-09-190-965-1
; Sequence 1, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-190-965-1

Alignment Scores:
Pred. No.: 3,89e-137 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: 3 Gaps: 3

US-10-089-688-1 (1-1053) x US-09-190-965-1 (1-337)
Qy 1 ATGCCGTTCCCGTTTGGGAAGTCTCACAAATCTCCAGACAGACATTGTGAAGAAATCTGAAG 60
Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
Qy 61 GAGACATGCTGTCTCGAAAAGCAAGACATTTCTCATATAAAAGCAGAAAAGCGGTACA 120
Db 23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
Qy 121 GAAGAAGTTTCCAAAATCTGGTCCCATGAAGAAATTTCTGTATGCGCACAAATGAAAA 180
```

```
Db 40 GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
Qy 181 GAGCCTCAGACAGACAGAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuTyrSerSerGlyLeuLeu 79
Qy 241 AGCACCTGTAGCTGATTTACAGCTCATTGACCTTTGAGGGCAAAAAGACGGCTCAA 300
Db 80 ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
Qy 301 ATTTTCAACCAATATTTCTCAGAACACAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
Db 100 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
Qy 361 TGCACCCCAACAGAAATATTTTGTTCATGTTTAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 120 SerAlaHisProHisIleLeuPheMetLeuLysGlyTyrGluAlaProGlnIleAla 139
Qy 421 CTAAATTGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAGAAATCATT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
Qy 481 TTGTGTCGGAACAGATTTTATGATTTTCTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 160 LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
Qy 541 GCTTCAGATGCTATTTGCCACATTTCAAGGATTTACTTACAGACATATAAATTTGCTCAGTGA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
Qy 601 GAATTTTGGACACAGCATTTATGATATGATATTTTTCAGTGAATATGAGAAGTTACTTCATCA 660
Db 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuGlnSer 219
Qy 661 GAAATTTATGTGACAAAAGACAGTCACTGGAAGCTTCTCGGTCAACTACTACTAGATAGA 720
Db 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
Qy 721 CACAACTTCACAATTATGACAAAATACATCAGTAAACCTGAGAACCTTCAAAATTAATGATG 780
Db 240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
Qy 781 AACCTGCTGGGAGACAAAAGTCCCAATCCAGTTTTCAGGCTTTCACGTTTTTAAGGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
Qy 841 TTTGTAGCCAAATCTTAAACAGACGAGCCCATCTAGACATCTCTCTCAAGAACACGAGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLysAsnGlnPro 299
Qy 901 AAACCTCATAGAGTTCTCTCAGCAAGTTTTCAGAACACGAGGACGAGTGTATGAGCAGTTCC 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
Qy 961 GTACCGACGAGAAATTTCCGGGTGCGATTTCAGGCTTAAACCGCGAGCGGTGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
Qy 1021 GATTTCAAGACAGACAGCT 1038
Db 331 AspLeuLysLysThrAla 336

RESULT 4
US-09-470-253-1
; Sequence 1, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
```



```
RESULT 7
US-09-190-965-5
; Sequence 5, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-09-190-965-5

Alignment Scores:
Pred. No.: 9.06e-110 Length: 377
Score: 1054.50 Matches: 205
Percent Similarity: 77.44% Conservative: 49
Best Local Similarity: 62.50% Mismatches: 61
Query Match: 58.49% Indels: 13
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-190-965-5 (1-377)

QY 1 ATGCCGTTCCCGTTGGAGTCTCAAAATCTCCAGCAGACATTTGCAAGATCTGAAG 60
Db 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20
QY 61 GAGACATGGCTGTTCTGAAAG-----CAAGAC 90
Db 21 GluValLeuThrIleLeuAspLysLeuProProLysLeuAspLysAspGlyAsnIle 40
QY 91 ATTTCTGATAAAAGCAGAAAGCTACAGAAGAGTTTCCAAAATCTGGTTCGCATG 150
Db 41 GlnSerAspLysLysTyrAspLysAlaLeuAspGluValSerLysAsnValAlaMetIle 60
QY 151 AAGAAATCTGTATGGCACAATGAAAGAGCTCAGACAGAA-----GCAGTA 201
Db 61 LysSerPheIleTyrGlyAsnAspSerAlaGluProSerSerGluHisValValGlnVal 80
QY 202 GCTCAACTGCTCAGAACTCTATAATAGTGGCTCCTTAGCACCTGCTAGCTGATTTA 261
Db 81 AlaGlnLeuAlaGlnGluValTyrAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100
QY 262 CAGCTCATTTGAGTGGGCAAAAGACGTTGCTCAAAATTTCAACAATTTCTCAGA 321
Db 101 ProLysPheGluPheGluCysLysLysAspValGlyGlnIlePheAsnLeuLeuArg 120
QY 322 AGACAAATGGPACGAGAACTCTACTGTTGAATACATCTGACCCCAACAGAAATTTTG 381
Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyrLeuGlyAlaArgProGluIleLeu 140
QY 382 TTCATGTTATTAAGAGGTGAATCTCCAGAAATAGCTCTAAATTTGGGAATATGTTA 441
Db 141 IleGlnLeuValGlnGlyTyrSerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160
QY 442 AGAAGATGTCATGAGATCAACACTTGCAGAAATCATTTTGTGTCGGAACAGATTTAT 501
Db 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTyrSerAspValPheTyr 180
QY 502 GATTTCCTCAGATATGTCGAATATGTCACATTTTGACATAGCTTCAGATGCTTCCACA 561
Db 181 ThrPhePheLeuTyrValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200

US-09-470-253-5
; Sequence 5, Application US/09470253
; Patent No. 635371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-09-470-253-5

Alignment Scores:
Pred. No.: 9.06e-110 Length: 377
Score: 1054.50 Matches: 205
Percent Similarity: 77.44% Conservative: 49
Best Local Similarity: 62.50% Mismatches: 61
Query Match: 58.49% Indels: 13
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-470-253-5 (1-377)

QY 1 ATGCCGTTCCCGTTGGAGTCTCAAAATCTCCAGCAGACATTTGCAAGATCTGAAG 60
Db 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20
QY 61 GAGACATGGCTGTTCTGAAAG-----CAAGAC 90
Db 21 GluValLeuThrIleLeuAspLysLeuProProLysLeuAspLysAspGlyAsnIle 40
QY 91 ATTTCTGATAAAAGCAGAAAGCTACAGAAGAGTTTCCAAAATCTGGTTCGCATG 150
Db 41 GlnSerAspLysLysTyrAspLysAlaLeuAspGluValSerLysAsnValAlaMetIle 60
QY 151 AAGAAATCTGTATGGCACAATGAAAGAGCTCAGACAGAA-----GCAGTA 201
Db 61 LysSerPheIleTyrGlyAsnAspSerAlaGluProSerSerGluHisValValGlnVal 80
QY 202 GCTCAACTGCTCAGAACTCTATAATAGTGGCTCCTTAGCACCTGCTAGCTGATTTA 261
Db 81 AlaGlnLeuAlaGlnGluValTyrAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100
QY 262 CAGCTCATTTGAGTGGGCAAAAGACGTTGCTCAAAATTTCAACAATTTCTCAGA 321
Db 101 ProLysPheGluPheGluCysLysLysAspValGlyGlnIlePheAsnLeuLeuArg 120
QY 322 AGACAAATGGPACGAGAACTCTACTGTTGAATACATCTGACCCCAACAGAAATTTTG 381
Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyrLeuGlyAlaArgProGluIleLeu 140
QY 382 TTCATGTTATTAAGAGGTGAATCTCCAGAAATAGCTCTAAATTTGGGAATATGTTA 441
Db 141 IleGlnLeuValGlnGlyTyrSerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160
QY 442 AGAAGATGTCATGAGATCAACACTTGCAGAAATCATTTTGTGTCGGAACAGATTTAT 501
Db 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTyrSerAspValPheTyr 180
QY 502 GATTTCCTCAGATATGTCGAATATGTCACATTTTGACATAGCTTCAGATGCTTCCACA 561
Db 181 ThrPhePheLeuTyrValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200
```

RESULT 10	
US-09-248-796A-15202	
Sequence 15202, Application US/09248796A	
Patent No. 6747137	
GENERAL INFORMATION:	
APPLICANT: Keith Weinstock et al	
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA	
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS	
FILE REFERENCE: 107196.132	
CURRENT APPLICATION NUMBER: US/09/248,796A	
CURRENT FILING DATE: 1999-02-12	
PRIOR APPLICATION NUMBER: US 60/074,725	
PRIOR FILING DATE: 1998-02-13	
PRIOR APPLICATION NUMBER: US 60/096,409	
PRIOR FILING DATE: 1998-08-13	
NUMBER OF SEQ ID NOS: 28208	
SEQ ID NO 15202	
LENGTH: 125	
TYPE: PRT	
ORGANISM: Candida albicans	
US-09-248-796A-15202	
Alignment Scores:	
Pred. No.:	5.15e-10
	Length: 125

```

; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15202
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15202

Alignment Scores:
Pred. No.: 5,15e-10
Length: 125

```


Db 264 AspGluLeuLysGluLysThrLysGlnPheAspAep-----SerLysLysLys 279
Qy 529 ACATTGACATAGCTTCAGATGCATTTGCCACATTCAGGATTTACTTACAGACATATAA 588
Db 280 LeuThrGluLeuGluAsnAspLeuThrSerThrLysLysGluLeuGluThr---GluLys 298
Qy 589 TTGCTCAGTCAGCAATTT-----TTGAAACAGCATTTATGATAGATTTTTCAGTGAATAT 642
Db 299 ThrGlnThrSerLysPheLysAsnLeuGluGluArgLysAspLysGluLeuValLysLeu 318
Qy 643 GAGAAG-----TTACTTCATTCAAGAAATATTGTACAAAAGACAGCTACTCAAG 693
Db 319 AsnLysGluLeuGluLeuLysAsnAspAsnSerGlyAlaLysLysGluLeuGlu 338
Qy 694 CTTCTCGGTGAACTACTACTAGATAGACACAACCTTCACAAATATGACAAAATACATCAGT 753
Db 339 LysValSerLysLeu-----GluSerGluLeuGluLeuSerLysGluLeuGlu 355
Qy 754 AAACCTGAGAACCTCAAAATTAATGATGAACCTGCTGCGAGACAAAGTCGCAACATCCAG 813
Db 356 AspLysLysSerValMetLysGlnHisAspGluLeuLysGluGlnThrLysGluLysAsn 375
Qy 814 TTTGAGGCTTTTCAGCTTTTAAAGGTGTTTGTAGCCAACTCCTAAACAGACGACGCCCATC 873
Db 376 GlnGluLeuGlnLysValThrLysAspTyr-----SerThrThrLysLysLys 391
Qy 874 CTAGACATCTCTCAAGAACCGCCMAACTCATAGATTCCTCAGCAAGTTTCAG--- 930
Db 392 LeuAspGluLeuGlnLysGluLeuAspAlaAlaLeuSerPheLysAspLysPheGluThr 411
Qy 931 --AACGACGAGCGGATTGTATGACGAGTTCCTGACGACGACGAGCAATTCGCGGTGAT 987
Db 412 AlaSerAlaLysLeuThrGlnSerThrSerAspLeuGluAlaAlaAsnLysLysLeuAsn 431
Qy 988 TTACGCGTTAAACCGCGACGCGTGGGATCAGGATTTGAAGACACCGAGCTCAGCAAGAA 1047
Db 432 IleLeuLysSerGluLysGluLysThrGluGlnGluLeuGluLysLeuThrLysGlnHis 451
Qy 1048 GCT 1050
Db 452 Ala 452

RESULT 13
US-09-134-000C-5420
; Sequence 5420, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5420
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5420

Alignment Scores:
Pred. No.: 0.000946 Length: 541
Score: 113.00 Matches: 89
Percent Similarity: 38.76% Conservative: 61
Best Local Similarity: 23.00% Mismatches: 141
Query Match: 6.27% Indels: 96
DB: 4 Gaps: 21

US-10-089-688-1 (1-1053) x US-09-134-000C-5420 (1-541)

Qy 40 GACATTGCTGAAGATCTGAAGAGAGCATGCTGCTTTCTGGAAAAGCAA---GACATTTCT 96
Db 173 AspIleIleLys---LeuProPheAlaLeuLysMetLeuAspLysHisLeuGluIleSer 191
Qy 97 GAT-----AAAAAAGCAGAAAGGCTACAGAA 126
Db 192 AspGlyAspLeuProHisAspAlaIleValLeuSerArgGluLysLeuLysAsnSerGlu 211
Qy 127 GTTTCCAAAATCTGGTTGCCATGAAGAAATCTGTATGGCACAATAAGAAAAAGAGCCT 186
Db 212 PheThrLysAsnAspTyrSerLeuLeuThrLeuTyrSerHisAlaAsnProLysLys--- 230
Qy 187 CAGACAGACAGCTAGCTCAACTGCTCAAGAA-----CTCTATATATAGTGGGCTCCTT 240
Db 231 -----GlnLeuAlaLeuGluValSerIlePheAsnSerTyrSerThr 244
Qy 241 AGCACCCCTG-----GTAGCTGATTTTACAGCTCATTCGAC 273
Db 245 AsnCysLeuLeuLeuValTyrLeuGlnAspGluLysValLysGluThrAspLeuValLe 264
Qy 274 TTTGAGGGCAAAAAGACGCTGGCTCAA---ATTTCAC-----AATATTCTCAGAAGA 324
Db 265 IleGluAsnAlaIleAspValLeuGlnLysPheAsnThrGluAsnLeuLeuLysLys 284
Qy 325 CAAATTGGTACGAGAACTCCTCTGTTGAATACATCTGCACCCACAGAAATATTTTGTTC 384
Db 285 GluArgTyrThrArgLeu-----AsnAsnLeuAlaAsp 295
Qy 385 ATGTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGCAATTAATGTTAAGA 444
Db 296 AlaIleLeuGln-----AsnThrProGlnAsnProAspGluLeuAsnSerLeuLeuArg 313
Qy 445 GAATGATCATCAGATCAACCACTTGCAAAAATCATTTTGTGTCGGAACAGTTTATGAT 504
Db 314 GluValGlnMetGlnGlu-----LeuAsp 321
Qy 505 TTCTTCAGATATCTCGAATGTCAACATTTGACATAGCTTCAGATGATTCGCACATTC 564
Db 322 SerTyrGlnAlaIleAlaPheSerThrLysAspThrAsnThrGlnLeuMetLysGluArg 341
Qy 565 AAGGATTTACTTACAAGACATAAATTTGCTCAGTGCAGAAATTTTGGAAACAG---CATTAT 621
Db 342 IleIleAlaLeuLeuArgThrLeuArgValLysSerIlePhePheAspGlnLeuAsnTyr 361
Qy 622 GATAGATTTTTCAGTGAATATGAGAAGTACTTTCATTTCAGAAAATATATGTGCAAAAAGA 681
Db 362 SerAlaValLeuPheAsnPheAsnGlu-----SerAspGlyLysIleThrLysLeu 378
Qy 682 CAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGACACAACTTCACAATATGACA 741
Db 379 GlnLeuSerArgLeuLeuAlaGluLeuLeuGluAsnAspThrLeuThrValAlaVal 398
Qy 742 AATACATCATGATAAACCCTGAGAACCTCAAAATTAATG-----ATGAACCTGCTG 789
Db 399 SerSerLeuLysSerArgGluGlyIleLysGluLeuLeuLysCysLeuAspIleLeu 418
Qy 790 CGAGACAAAAGTCGCAACATCCAGTTTGGAGCCCTTTCAC-----GluThrPheTyrAsnGlyProIleValThrLeu 432
Db 419 ArgPheAsn-----GluThrPheTyrAsnGlyProIleValThrLeu 432
Qy 829 -----GTTTTTAAGGTGTTGTA-----GCCAATCTCAACAAG 861
Db 433 AlaAspIleGlyValPheLysAsnPheIleArgGluAspGlnLeuGluLeuAspGlu 452
Qy 862 ACGCAGCCC---ATCCTAGACATCTCCTCCTCAGAACACAGGCCCAAACTCATAGATTCCTC 918
Db 453 LeuIleProArgAlaLeuTyrGlnLeuAlaGluAsnAsnTyrAspLeuPheGluThrLeu 472
Qy 919 ---AGCAAGTTTTCAGAAACGACGACGGATTTGATGAGCAGATTCCTCCGTCACGACGCAAT 975
Db 473 TyrSerPhePheGlnAsnAsnArgAsnTyrLysGlnThrSerGluAlaMetPheLeuHis 492
Qy 976 TCCCGGGTCGATTTACGCGTTAAACCGCGGACGCGTGGGATCAGG----- 1020

```
Db 493 SerLysThr-----IIeArgTyrArgLeuAsnLysValGluGlnLeuLeuAspIle 509
QY 1021 GATTGTAAGAGACGACGCTCAG 1041
Db 510 AspLeuAlaAsnProLeuGln 516
RESULT 14
US-09-417-485D-6
; Sequence 6, Application US/09417485D
; Patent No. 6541202
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; FILE REFERENCE: 47714-5009-US
; CURRENT APPLICATION NUMBER: US/09/417,485D
; CURRENT FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2184
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (330)..(335)
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
US-09-417-485D-6
Alignment Scores:
Pred. No.: 0.00183 Length: 2184
Score: 113.00 Matches: 68
Percent Similarity: 40.94% Conservative: 54
Best Local Similarity: 22.82% Mismatches: 98
Query Match: 6.27% Indels: 78
DB: 4 Gaps: 18
US-10-089-688-1 (1-1053) x US-09-417-485D-6 (1-2184)
QY 265 CTCATTGACTTGGGCAAAAAGACGTGCTCAAAATTTTCAACAAT-----312
Db 401 LeuLeuAsnTyrIlePheLysSerAspGlnThrTyrPheTyrHisAsnAsnPheIleAsp 420
QY 313 -----ATTCTCAGAAGACAAAT-----GGTAGCAGAACTCTCTACTGTTGAA 354
Db 421 GluTyrLysGlnLysIleCysLysGlnIleLysCysSerThrLysLysAsnAspIleSer 440
QY 355 TACATCTGCACC-----CAACAGAAATTTTGTTCATGTTATTGAAA-----GGG 399
Db 441 HisIleIleThrSerArgLysGluAsnHisLeuPheHisValGlnLysLeuGluAsnAsn 460
QY 400 TATGAATCTCCAGAATAGCTTAATTTGGAATAATGTTAAGAGATGATCAGACAT 459
Db 461 TyrLysHisProAsnIle-----AsnLysGlnLeuArgLys 472
QY 460 GAACCACTTGCAAAAATCATTTTGTGTCGGAACACAGTTTATGATTTCTTCAGATATGTC 519
Db 473 ThrLysIleLeuLysTyrValTyr-----AsnTyrPheLysGluPheIleAsnVal 490
QY 520 GAAATGTCAACATTTGACATAGCTTCAGATGCAATTTGCCACATTCACAGATTACTTACA 579
Db 491 IleAsnThrLysPheGlyLysIleTyrArgLysPhePheProArgLysHisIleLeuAsn 510
QY 580 AGA---CATAATG-----591
Db 511 LysIleHisIlePheLysIleIleArgLeuGlnIleIleLysLysTyrArgIleIle 530
QY 592 -----CTCAGTGCAGAAATTTTGGAAACAGCAT---TATGATAGATTTTTCAGTGAA 639
Db 531 AsnIleArgMetAsnArgLysPheIleLysGlnLysValTyrAspThrPhePheLysAsn 550
```

```
QY 640 TATGAG-----AAGTTACTTCAATTCAGAAATATTATG 672
Db 551 TyrAspPheLeuSerPheSerPheLysThrTyrLysIleIleAsnPheMetValTyrIle 570
QY 673 ACAAAGA---CAGTCACTGAAGTTCGCGTGAACACTACTAGATAGACACAACCTTC 729
Db 571 ThrLysLysCysIleProIleLysLeuLeuGly-----SerLysHisAsnPhe 586
QY 730 ACATTTATGACAAAATACATCATGATAA-----CCTCAGAACCTCAAA 771
Db 587 LysIlePheLeuLysAsnValLysLysPheLeuLeuPheAsnTyrLysGluSerPheSer 606
QY 772 TTAATGATGAACCTGCTGCGAGACAAAGTCGCAACATCCAGTTTCAGGCTTCCTCAGTT 831
Db 607 LeuAsnGlnValMetLysAsnIleLysValLysAsnIle---PheGlnLys---LysIle 624
QY 832 TTTAAGGTGTTTGTAGCCAATCTCTAACAAGACGACGCCATCTCTAGAC-----879
Db 625 SerLysTyrAsnIleLysAsnArgIleLeuLeuLysAsnIlePheAspAsnAsnTyrGlu 644
QY 880 -----ATCTCTCTCAAGAACCGACGCAAACTCATAGAGTTCTCTCAGC-----AAG 924
Db 645 AsnLysIleLeuHisArgAsnAsnLysGluIleIleThrAsnIleAsnAspAsnIleLys 664
QY 925 TTTCAGAAGCAGACGACGAGGATTGTATGACGAGCTTC-----GTACCGACGACG 972
Db 665 IleTyrAsnLysLysAsnAspAsnLeuAsnAsnSerPheLysIleLysThrThr 682
RESULT 15
US-09-134-001C-4053
; Sequence 4053, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4053
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4053
Alignment Scores:
Pred. No.: 0.0018 Length: 533
Score: 110.50 Matches: 73
Percent Similarity: 40.00% Conservative: 57
Best Local Similarity: 22.46% Mismatches: 108
Query Match: 6.13% Indels: 87
DB: 3 Gaps: 18
US-10-089-688-1 (1-1053) x US-09-134-001C-4053 (1-533)
QY 40 GACATTGTGAAGAACTCTGAGGAGACGATGCTGTTCTTG-----GAAAGCAA 87
Db 237 AspIleLeuArgGlnPheSerAspSerIleAspValLeuIleThrGlyHisGlnHisArg 256
QY 88 GACATTTCTGATAAAGCAGCAAAAGCT---ACAGNAGAAAGTTTCCAAAATCTGGTT 144
Db 257 AspIleAlaThrIleLysAsnGlnThrAlaIleIleGlnProGlySerLysGlyThrLys 276
QY 145 GCCATGAAGAAATCTGTATGCGCAAAATCAAAA-----180
Db 277 ValGlyLysIleValIleGluTyrThrHisAspLysLysValLeuLysGluCysAsn 296
```

```
Qy 181 -----GAGCCTCAGACAGACAGCAGTAGCTCAACTT 210
Db 297 LeuMetAsnValHisAsnSerThrSerPheLysProAsnAspGluAspIleAala---Leu 315
Qy 211 GCTCAAGAACCTCTATATAGTGGCTCCTTAGCACCCCTGGTAGCTGATTTACAGCTCATT 270
Db 316 ArgAsnGlnLeu-----GluAspTrpLeuAspThrGlnIleAlaGluLeuProTyrAla 333
Qy 271 -----GACTTTGAGGGCAAAAAGAC-----GTGGCTCAAAATTTTC 306
Db 334 MetArgIleAsnAsnSerPheGluAlaArgLysSerProHisAlaPheValAsnLeuLeu 353
Qy 307 AACATATATTCAGAGACAAATTTGGTAGAGAACTCCTACTGTGTGAATACATCTGCACC 366
Db 354 AsnTyrIleLeuLeuGluLysSerGly-----AlaAspIleAlaCysThr 368
Qy 367 CAACAGAAATATTTTGTTCATCTTATGAAGGGTATGAATCTCCAGAAATAGCTCTAAAT 426
Db 369 -----AlaLeuPheAspSerAlaAsnGlyPheAspGluLys----- 380
Qy 427 TGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAGAAATCAATTTGTGG 486
Db 381 -----ValThrMetArgAspIleIleAsnAsnTyrPro----- 391
Qy 487 TCGGAACAGTTTATGATTTCTTCAGATATCTCGAAATGTCACATTTGACATA----- 540
Db 392 -----PheProAsnThrPheLysValIleGluLeuSerGlyLysAspIleLysLeu 408
Qy 541 GCTTCAGATGCATTTGCCACATTCACAGGATTTACTTACAGACATAAATTCCTCAGTGCA 600
Db 409 AlaIleGluArgSerAlaSerTyrPheAspIleValAsnHisLysIleThrValAsnLys 428
Qy 601 GAATTTTGTGAA-----CAGCATTTATGAT---AGATTTTTCAGTGAATATGAGAAG 648
Db 429 GluPheLeuGluProLysProGlnHisPheAsnTyrAspIlePheAlaGlyIleGlnTyr 448
Qy 649 TTACTTCATTCAGAAAATATGTGACAAAAGACAGTCACCTGAAGCTTCTCGGTGAACATA 708
Db 449 ThrIleHisValSerHisProTyrGlyGluArgValSerAspLeuLeuIleAsnAspAla 468
Qy 709 CTACTAGATAGACACAAC---TTCACAATTTATGACAAA----- 744
Db 469 ProLeuGlnSerAspGlnIleTyrThrIleCysValAsnAsnTyrArgAlaValGlyGly 488
Qy 745 -----TACATCAGTAAACCT-----GAGAACCTC 768
Db 489 GlyAsnTyrAspMetTyrValAsnLysProValIleLysAspIleGlnIleGluGlyAla 508
Qy 769 AATTAATGATGAACCTGCTG-----CGAGACAAAAGTCGCAACATCCAGTTTGAGGCC 822
Db 509 GlnLeuLeuIleAspTyrLeuSerHisAsnAspLeuSerGlnIleProGlnValIleAsp 528
Qy 823 TTTCACGCTTTTTRAG 837
Db 529 PheAsnValValLys 533
```

Search completed: November 10, 2004, 20:54:02
Job time : 47.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:43:16 ; Search time 165.5 Seconds
(without alignments)
4494.826 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atccgttcctccgtttgggaa.....cagctcagcagaagcttaa 1053

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 3133240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10089688/runat 10112004 152249 4931/app query.fasta_1.1223
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOPEXT=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USP=US10089688 @CGN 1 1 199 @runat 10112004 152249 4931
-NCPU=6 -ICPU=3 -NO MAP= -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

* Query
Result

No.	Score	Match	Length	DB	ID	Description
1	1635.5	90.7	341	16	US-10-408-765A-911	Sequence 911, App
2	1635.5	90.7	496	14	US-10-239-079-5	Sequence 5, Appli
3	1635.5	90.7	552	14	US-10-239-079-6	Sequence 6, Appli
4	1614.5	89.5	341	14	US-10-025-730-3	Sequence 3, Appli
5	1297.5	72.0	337	14	US-10-025-730-1	Sequence 1, Appli
6	1117	62.0	339	14	US-10-025-730-4	Sequence 4, Appli
7	1054.5	58.5	377	14	US-10-025-730-5	Sequence 5, Appli
8	750	41.6	446	15	US-10-425-114-52177	Sequence 52177, A
9	747.5	41.5	337	17	US-10-425-114-322057	Sequence 322057,
10	747	41.4	336	16	US-10-437-963-161075	Sequence 161075,
11	741	41.1	412	15	US-10-425-114-54669	Sequence 54669, A
12	721.5	40.0	339	15	US-10-424-599-224185	Sequence 224185,
13	652	36.2	336	17	US-10-425-115-201848	Sequence 201848,
14	649	36.0	351	15	US-10-425-114-46025	Sequence 46025, A
15	639.5	35.5	337	16	US-10-437-963-183779	Sequence 183779,
16	635.5	35.2	342	15	US-10-425-114-48872	Sequence 48872, A
17	635.5	35.2	363	17	US-10-425-115-215481	Sequence 215481,
18	593	32.9	334	16	US-10-437-963-146612	Sequence 146612,
19	501	27.8	337	15	US-10-424-599-209662	Sequence 209662,
20	459.5	25.5	256	17	US-10-425-115-322099	Sequence 322099,
21	400.5	22.2	197	16	US-10-767-701-37391	Sequence 37391, A
22	387	21.5	274	17	US-10-425-115-233795	Sequence 233795,
23	379.5	21.0	173	15	US-10-424-599-158029	Sequence 158029,
24	362	20.1	205	17	US-10-425-115-201847	Sequence 201847,
25	353.5	19.6	140	16	US-10-767-701-42587	Sequence 42587, A
26	322	17.9	351	17	US-10-425-115-201850	Sequence 201850,
27	321	17.8	184	15	US-10-424-599-149595	Sequence 149595,
28	286.5	15.9	294	16	US-10-437-963-119830	Sequence 119830,
29	255	14.1	149	17	US-10-425-115-350620	Sequence 350620,
30	255	14.1	161	15	US-10-425-114-70029	Sequence 70029, A
31	218	12.1	90	17	US-10-425-115-358599	Sequence 358599,
32	209	11.6	135	11	US-09-864-408A-6108	Sequence 6108, Ap
33	178	9.9	94	17	US-10-425-115-282537	Sequence 282537,
34	168	9.3	83	15	US-10-425-114-44450	Sequence 44450, A
35	135.5	7.5	116	15	US-10-424-599-210348	Sequence 210348,
36	126	7.0	1182	15	US-10-282-122A-53445	Sequence 53445, A
37	125.5	7.0	74	15	US-10-425-114-56749	Sequence 56749, A
38	116.5	6.5	888	10	US-09-893-519A-73	Sequence 73, Appli
39	113	6.3	2184	14	US-10-304-095-6	Sequence 6, Appli
40	111.5	6.2	70	15	US-10-424-599-210351	Sequence 210351,
41	111.5	6.2	790	16	US-10-408-765A-2178	Sequence 2178, Ap
42	110.5	6.1	797	16	US-10-437-963-163096	Sequence 163096,
43	110.5	6.1	4131	14	US-10-369-493-5136	Sequence 5136, Ap
44	109.5	6.1	367	15	US-10-282-122A-45776	Sequence 45776, A
45	109.5	6.1	564	15	US-10-282-122A-61170	Sequence 61170, A

ALIGNMENTS

RESULT 1
US-10-408-765A-911
; Sequence 911, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 66088.465
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 911
; LENGTH: 341
; TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-408-765A-911

Alignment Scores:
Pred. No.: 3,18e-151 Length: 341
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 16 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-408-765A-911 (1-341)
QY 1 ATGCGCTCCGTTGGAGTCTCACAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAGCAAGAAAGGCTACA 120
DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAAGTTTCCAAAATCTGGTGGCCATGGAAGAAATCTGTATGCGCACAAATGAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCTCAGCAGACAGTAGTCACTCACTGCTCAAGAACTCTATAATAGTGGCTCCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTAGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGAGCGTCAAA 300
DB 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACATATTTCTCAGAAAGCAATTTGGTACGAGAACTCTACTGTTGAATACATC 360
DB 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCCAACAGATATTTTGTTCATGTTATTTGAAGGTTATGAATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAATTTGGAATTAATTTAAGAGATGCAATGCAATGACATGAACCACTTGCAAAATCATT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGCGTCGGAACAGTTTATGATTTCTCAGATATGTCGAATGTCGAATCTCAACATTTGACATA 540
DB 161 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCAATTTGCCACATTTCAAGGATTTACTTACAAGACATAAAATTTGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAATTTTGGACACAGCATATGATAGATTTTTCAGTGAATATGAGAAATTTCTTCAATTCAC 660
DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
QY 661 GAAATTTATGTCACAAAAGACAGTCACTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
QY 721 CACAACCTTCACAAATATGACAAAATACATCAGTAAACCTTGAAACCTCAAAATTAATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACCTGCTCGGAGACAAAGTCGCAACATTCAGTTTGGAGCCCTTTCACGTTTAAAGGTG 840
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCAATCTTAAACAGACGAGCCCATCTCTAGACATCTCTCAAGAACACGAGCC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300

RESULT 2
US-10-239-079-5
; Sequence 5, Application US/10239079
; Publication No. US20030148446A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP-1-ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gal4-ANIC-BP-1
; OTHER INFORMATION: fusion protein
US-10-239-079-5

Alignment Scores:
Pred. No.: 3,52e-151 Length: 496
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 16 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-239-079-5 (1-496)
QY 1 ATGCGCTCCGTTGGAGTCTCACAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
DB 156 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 175
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAGCAAGAAAGGCTACA 120
DB 176 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 195
QY 121 GAAGAAGTTTCCAAAATCTGGTGGCCATGGAAGAAATTTCTGTATGCGCACAAATGAAAA 180
DB 196 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 215
QY 181 GAGCTCAGCAGACAGTAGCTCACTGCTCAAGAACTCTATATAGTGGCTCCTT 240
DB 216 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 235
QY 241 AGCACCTGGTAGCTGATTTTACAGCTCATTGACTTTTCAGGGCAAAAAGAGCTGGCTCAA 300
DB 236 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 255
QY 301 ATTTTCAACATATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTACTGTTGAATACATC 360
DB 256 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 275
QY 361 TGCACCCCAACAGATATTTTGTTCATGTTATTTGAAGGTTATGAATCTCCAGAAATAGCT 420
DB 276 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 295
QY 421 CTAATTTGGAATTAATTTAAGAGATGCAATGCAATGACATGAACCACTTGCAAAATCATT 480
```

Db 296 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 315
Qy 481 TTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 316 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 335
Qy 541 GCTTCAGATGCTATTCACACATTCACAGATTTACTTACAGACATAAATGCTCAGTGCA 600
Db 336 AlasrAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 355
Qy 601 GAAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAAGTTACTTCATTCA 660
Db 356 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 375
Qy 661 GAAATTTATGTGACAAAAGACAGTCACTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db 376 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 395
Qy 721 CACAACTTCACAAATATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATCATG 780
Db 396 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 415
Qy 781 AACCTGCTGCAGACAAAAGTCGCAACATCCAGTTTGAGGCTTTTCACGTTTAAAGGTG 840
Db 416 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 435
Qy 841 TTTGTAGCAATCTTAACAGACGAGCCCATCTAGACATCTCTCAAGAACACGAGCC 900
Db 436 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 455
Qy 901 AAACCTAGAGTTCCTCAGCAAGTTTCAGAACACAGGCGGATTTGTATGAGCAGTTCC 960
Db 456 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu-----IleArg 470
Qy 961 GTACCGACGCAATTCGCGGTGATTTACGCGTTAAACCGCGGCGCGGTGATCAGG 1020
Db 471 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 486
Qy 1021 GATTGGAAGACAGCTCAGCAAGAGCT 1050
Db 487 AspLeuLysArgProAlaGlnGlnGluAla 496

RESULT 3

US-10-239-079-6
; Sequence 6, Application US/10239079
; Publication No. US2003014846A1
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP-1-ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LexA-ANIC-BP-1
US-10-239-079-6

Alignment Scores:

Pred. No.:	3,62e-151	Length:	552
Score:	1635.50	Matches:	329
Percent Similarity:	94.29%	Conservative:	1
Best Local Similarity:	94.00%	Mismatches:	11
Query Match:	90.71%	Indels:	9
DB:	14	Gaps:	2

US-10-089-688-1 (1-1053) x US-10-239-079-6 (1-552)

RESULT 4

Qy 1 ATGCCGTTCCGCTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAAATCTGAAG 60
Db 212 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 231
Qy 61 GAGAGCATGCTGCTTCTGGAAGAACAGACATTCTTGATAAAAGAGCAGAAAAGGCTACA 120
Db 232 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 251
Qy 121 GAGAGAGTTTCCAAAAATCTGGTGGCATGAAAGAAATCTGTATGGCACAATAATGAANA 180
Db 252 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 271
Qy 181 GAGCCTCAGACAGACAGCTAGCTCACTTGGCTCAAGAACTCTATATAGTGGCTCCTT 240
Db 272 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 291
Qy 241 AGCACCTCGTGTAGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGACGTGGCTCAA 300
Db 292 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 311
Qy 301 ATTTTCAACAATATCTCAGAAAGCAAAATTTGGTACGAGAACTCTCTACTGTGTGAATCATC 360
Db 312 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 331
Qy 361 TGCACCCACAGATATTTTGTTCATGTTATTCAGAGGTATGAATCTCCAGAAATAGCT 420
Db 332 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 351
Qy 421 CTAAATTTGGAATAATGTTAAAGAGAAATGTCATCAGACATCAACACACTTGCAAAATCATT 480
Db 352 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 371
Qy 481 TTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 372 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 391
Qy 541 GCTTCAGATGCTATTCGCCACATTCAGGATTTACTTACAAGACATAAATGCTCAGTGCA 600
Db 392 AlasrAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 411
Qy 601 GAAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAAGTTACTTCATTCA 660
Db 412 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 431
Qy 661 GAAATTTATGTGACAAAAGACAGTCACTGAACTTCTCGGTGAACTACTACTAGATAGA 720
Db 432 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 451
Qy 721 CACAACTTCACAATATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
Db 452 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 471
Qy 781 AACCTGCTGCAGACAAAAGTCCAACTCAGTTTGGGCTTTTACAGCTTTTAAAGGTG 840
Db 472 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 491
Qy 841 TTTGTAGCCCAATCTTAACAGACGCGAGCCCATCTAGACATCTCTCAAGAACACGAGCC 900
Db 492 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 511
Qy 901 AAACCTATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGAGCAGGATTTGTATGAGCAGTTCC 960
Db 512 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu-----IleArg 526
Qy 961 GTACCGACGAGAAATTCGCGGTGATTTACGCGTTAAACCGCGGCGCGGTGATCAGG 1020
Db 527 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 542
Qy 1021 GATTGGAAGACAGCTCAGCAAGAGCT 1050
Db 543 AspLeuLysArgProAlaGlnGlnGluAla 552

```
US-10-025-730-3
; Sequence 3, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: 9262934
US-10-025-730-3

Alignment Scores:
Pred. No.: 3.64e-149 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 14 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-025-730-3 (1-341)

QY 1 ATGCGCTGCCGTTGGGAGTCTCACAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
Db 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTTCTGGAAGCAAGACATTTCTGATAAAGCAAGCAAGAGGCTACA 120
Db 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAGTTTCCAAAATCTGTTGCCATGAAGAAATTTCTGTATGGCACAAATGAAAAA 180
Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCTCTACAGACAGCAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGCTCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTAGCTGATTACAGCTATTGACTTTGAGGCGCAAAAGACGTGGCTCAA 300
Db 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACAATATTTCTCAGAGACAAATTTGTAGAGAACTCTCTACTGTTTGAATACATC 360
Db 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCACAGAAATATTTGTTCATGTTATTGAAAGGCTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAAATTTGGTAATATTTAAGAGAAATGTCATCAGACATGACCACTTGCAGAAATCATTT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGTCGTGGGACAGTTTTTATCATTTCTTCAGATATGTCGAATGTCAACATTTGACATA 540
Db 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCAATTTGCCACATTCAGGATTTTACTACAGACATATAATTTGCTCAGTGA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200

US-10-025-730-1
; Sequence 1, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-10-025-730-1

Alignment Scores:
Pred. No.: 4.22e-118 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: 14 Gaps: 3

US-10-089-688-1 (1-1053) x US-10-025-730-1 (1-337)

QY 1 ATGCGGTTCCCGTTTGGGAGTCTCACAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGTATAAAGCAAGACAGAAAGCTACA 120
```

```

Db      23 AspAsnLeuAlaLeuLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
QY      121 GAAGAGTTTCCAAATCTGTTGCCATGAAGAAATCTGTATGGCACAAATGAAAAA 180
Db      40 GluGluValSerLysSerLeuGlnAlaMetLysGluLeuLysGlyThrAsnGluLys 59
QY      181 GAGCCTCAGACAGAGTAGTACCTCAACTGCTCAAGAACTCTATATAGTGGCTCCTT 240
Db      60 GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuLysSerGlyLeuLeu 79
QY      241 AGCACCTGGTAGCTGATTACAGCTCATTTGAGTCTTGAAGGCAAAAAAGAGCTGGCTCAA 300
Db      80 ValThrLeuLeuAlaAspLeuGlnLeuLeuAspPheGluGlyLysLysAspValThrGln 99
QY      301 ATTTTCAACAATATCTCAGAAGCAAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
Db      100 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
QY      361 TGCACCCCAACAGATATTTTCTGATGTTTGAAGGGTATGAATCTCCAGAAATAGCT 420
Db      120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
QY      421 CTAAATTTGTGAATAATGTTAAGAGAAATGCAATCAGACATGAACCACTTGCAAAATCAT 480
Db      140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
QY      481 TTGTGGTGGGAACAGTTTATGATTCTTTCAGATATGTCGAATATGTCGAATTTGACATA 540
Db      160 LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
QY      541 GCTTCAGATGCTATGTCACATTCACAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
Db      180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY      601 GAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAATGTTACTTCAT 660
Db      200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
QY      661 GAAATTTATGTGACAAAGACAGTCTACTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db      220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 239
QY      721 CACAATCTCACAATTATGACAAATATACATCAGTAAACCTCAGAACCTCAATTAATCATG 780
Db      240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY      781 AACCTGCTGCGAGACAAAGTCGCAACATCCAGTTTCAGGCTTTCACGTTTTTAAGGTG 840
Db      260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY      841 TTTGTAGCCAATCTTAACAAGACGCGCCATCCTAGACATCCTCTCAAGAACCCAGGCC 900
Db      280 PheValAlaSerProHisLysThrGlnProIleValGluLeuLeuLysAsnGlnPro 299
QY      901 AAATCTCATAGTTCCTCAGACAGTTTCAGAACGACGAGCGAGTGTATGACAGTTCC 960
Db      300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY      961 GTACCGACGACGAAATCCCGGGTCGATTTACGGCTTAAACCGCGAGCGGTGGATCAGG 1020
Db      315 -----AspGluGlnPheAlaAspGluLysAsnTyrIleuLysGlnIleArg 330
QY      1021 GATTGAAGACAGCAGCT 1038
Db      331 AspLeuLysLysThrAla 336

```

RESULT 6

US-10-025-730-4

```

; Sequence 4, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom

```

```

; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: -
; OTHER INFORMATION: g1794137
US-10-025-730-4

```

Alignment Scores:

```

Pred. No.: 2,07e-100 Length: 339
Score: 1117.00 Matches: 220
Percent Similarity: 84.95% Conservative: 51
Best Local Similarity: 68.97% Mismatches: 40
Query Match: 61.95% Indels: 8
DB: 14 Gaps: 5

```

US-10-089-688-1 (1-1053) x US-10-025-730-4 (1-339)

```

QY      1 ATCCGCTCCCGTTGGAGAGTCTCACAAATCTCCAGCACACATTGTGAAGATCTGAAG 60
Db      1 MetProLeu---PheGlyLysSerGlnLysSerProValGluLeuValLysSerLeuLys 19
QY      61 GAGAGCATGCTGCTGTTCTGGAAGAGCAAGACATTTCTGATAAAAGCAGAAAGCGGTACA 120
Db      20 GluAlaIleAsnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln 36
QY      121 GAAGAAGTTTCCAAAAATCTGGTGGCATGAAGAAATCTGTATGGCACAAATGAAAAA 180
Db      37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuHisGlySerSerAspAla 56
QY      181 GACCTCTCAGACAGAA---GCAGTAGCTCACTTCTCGTCAAGAACTCTATATAGTGGCTC 237
Db      57 GluProProAlaAspTyrValValAlaGlnLeuSerGlnGluLeuTyrAsnSerAsnLeu 76
QY      238 CTTAGCACCTGCTAGTCTGATTTACAGCTCATTTGAGCTTTTCAGGCGCAAAAAGAGCTGCT 297
Db      77 LeuLeuLeuLeuIleGlnAsnLeuHisArgIleAspPheGluGlyLysLysHisValAla 96
QY      298 CAAATTTTCAACAATATCTCAGACAGCAAAATTTGGTACGAGAACTCCTACTGTTGAATAC 357
Db      97 LeuIlePheAsnAsnLeuLeuArgGlnIleGlyThrArgSerProThrValGluTyr 116
QY      358 ATCTGCGACCCACAGCAATATTTTGTTCATGTTATTGAAGGGTATGATCT-----CCA 411
Db      117 IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlyTyrGluAspAlaHisPro 136
QY      412 GAAATAGCTCTAAATTTGCGAATAATGTTAAGAGAAATGCAATCAGACATGAACCACTTGCA 471
Db      137 GluIleAlaLeuAsnSerGlyThrMetLeuArgGluCysAlaArgTyrGluAlaLeuAla 156
QY      472 AAAATCATTTTGTGGTGGAAACAGTTTATGATTTTCTTCAGATATGTCGAAATGTCACAA 531
Db      157 LysIleMetLeuHisSerAspGluPhePheLysPhePheArgTyrValGluValSerThr 176
QY      532 TTTGACATAGCTTCAGATCATTGCGACATTTGCCATTCACAGGATTTACTTACAGACATTAAT 591
Db      177 PheAspIleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeu 196
QY      592 CTCAGTGCAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAA---TATGAGAAG 648
Db      197 LeuCysAlaGluPheLeuAspAlaAsnTyrAspLysPhePheSerGlnHisTyrGlnArg 216

```

7

US-10-425-114-52177

Alignment Scores:

Pred. No.: 2.06e-64 Length: 446
Score: 750.00 Matches: 147
Percent Similarity: 66.98% Conservative: 64
Best Local Similarity: 46.67% Mismatches: 100
Query Match: 41.60% Indels: 4
DB: 15 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-425-114-52177 (1-446)

```
Qy 13 TTTCGGAGTCTCACAATCTCCAGCAGCATCTGGAAGTCTGAAGGAGCATGGCT 72
Db 115 PhelysSerLysProArgThrProValAspIleValArgGlnThrArgGluCysLeuVal 134
Qy 73 GTTCTGGAAAGCAGACATTCTCT-----GATAAAGGAGCAGAAAGGCTACAGAA 123
Db 135 TyrLeuAspLeuHisSerAspSerArgSerGlyAspAlaLysArgGluGluLysMetThr 154
Qy 124 GAAGTTTCCAAAATCTGGTTCATGAAAGAAATCTGTATGGCACAATAATGAAAAAGAG 183
Db 155 GluLeuSerLysAsnIleArgAspMetLysSerIleLeuTyrGlyAsnGlyGluSerGlu 174
Qy 184 CCTCAGACAGACAGTAGCTCAACTGCTCAAGACTCTATATAGTGGCTCCTTAGC 243
Db 175 ProValThrGluAlaCysValGlnLeuThrGlnGluPhePheArgGluAsnThrLeuArg 194
Qy 244 ACCCTGTAGCTGATTTACAGCTCATTTGAGCTTTGAGGCGCAAAAGAGCTGCTCAAAAT 303
Db 195 LeuLeuIleHisLeuProLysLeuAsnLeuGluThrArgLysAspAlaThrGlnVal 214
Qy 304 TTCAACAATATCTCAGAAGCAAAATGGTACGAGAACTCCTACTGTGTAATACATCTGC 363
Db 215 ValAlaAsnLeuGlnArgGlnValSerSerLysIleValAlaSerGluTyrLeuGlu 234
Qy 364 ACCCAACAGATATTTTGTTCATGTTATTTGAAAGGTATGAATCTCCAGAAATAGCTCTA 423
Db 235 SerAsnLysAspLeuLeuAspIleLeuIleLeuGlyTyrGluAsnMetAspIleAlaLeu 254
Qy 424 AATGTGGAATATGTTAAAGAGATGATCAGACATCAACACCATTTGCAAAATCATTTTG 483
Db 255 HisTyrGlyAlaMetLeuArgGluCysIleArgHisGlnSerIleAlaArgTyrValLeu 274
Qy 484 TGGTCGGAACAGTTTATGATTTCTTCAGATATGCGAAATGCAATTTGACATAGCT 543
Db 275 GluSerGluHisMetLysLysPheAspTyrIleGlnLeuProAsnPheAspIleAla 294
Qy 544 TCAGATCATTGTCACATTCAGGATTTACTTACAAGACATAAATTTGCTCAGTGCAGAA 603
Db 295 SerAspAlaSerAlaThrPheLysGluLeuLeuThrArgHisLysAlaThrValAlaGlu 314
Qy 604 TTTTGGACAGCATTATGATAGATTTTTCAGTGAATAT--GAGAGATTACTTCATTCA 660
Db 315 PheLeuSerAsnAsnLysPhePhePheGluGluPheAsnSerArgLeuLeuSerSer 334
Qy 661 GAAATTTATGACAAAGACAGTCTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db 335 ThrAsnTyrIleThrLysArgGlnAlaIleLysLeuGluGlyAspMetLeuLeuAspArg 354
Qy 721 CACAACCTTCACATTTATGACAAATATACATCAGTMAACCTCAGAACCTCAATTAATGATG 780
Db 355 SerAsnAlaAlaValMetMetArgTyrValSerSerLysAspAsnLeuMetIleLeuMet 374
Qy 781 AACTGCTGAGACAAAAGTCGCAACATCCAGTTTTCAGGCGCTTTCACGTTTTTAAGGTG 840
Db 375 AsnLeuLeuArgAspSerSerLysAsnIleGlnIleGluAlaPheHisValPheLysLeu 394
Qy 841 TTTTGTAGCCCAATCTTAACAGAGCGCAGCCCATCTCTAGACATCTCTCCTCAAGAACCGGCC 900
Db 395 PheAlaAlaAsnLysAsnLysProGluValValAsnIleLeuValThrAsnArgSer 414
Qy 901 AAACATCATAGAGTTCTTCAGCAAGATTTTCAGAACGACGAGCGGAT 945
```

Db 415 LysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 429

RESULT 9

US-10-425-115-322057
Sequence 322057, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 322057
LENGTH: 337
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(337)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_5677C.1.pep
US-10-425-115-322057

Alignment Scores:

Pred. No.: 3.36e-64 Length: 337
Score: 747.50 Matches: 146
Percent Similarity: 66.46% Conservative: 64
Best Local Similarity: 46.20% Mismatches: 101
Query Match: 41.46% Indels: 5
DB: 17 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-425-115-322057 (1-337)

```
Qy 13 TTTCGGAGTCTCACAATCTCCAGCAGCATCTGGAAGTCTGAAGGAGCATGGCT 72
Db 5 PhelysSerLysProArgThrProValAspIleValArgGlnThrArgGluCysLeuVal 24
Qy 73 GTTCTGGAAAGCAGCA-----GACATTTCTGATAAAAAAGCAGAAAGGCTTACA 120
Db 25 TyrLeuAsp***SerThrArgThrProGlyAla**AspAlaLysArgGluGluLysMet 44
Qy 121 GAAGAAGTTTCCAAAATCTGGTTCATGAAAGAAATCTGTATGGCACAATAATGAAAAA 180
Db 45 ThrGluLeuSerLysAsnIleArgAspMetLysSerIleLeuTyrGlyAsnGlyGluSer 64
Qy 181 GAGCCTCAGACAGACAGTAGCTCAACTGCTCAAGACTCTATATAGTGGCTCCTT 240
Db 65 GluProValThrGluAlaCysValGlnLeuThrGlnGluPhePheArgGluAsnThrLeu 84
Qy 241 AGCACCTCTGTAGCTCATTTTACAGCTCATTTGAGGCGCAAAAGAGCTGGCTCAA 300
Db 85 ArgLeuLeuIleHisLeuProLysLeuAsnLeuGluThrArgLysAspAlaThrGln 104
Qy 301 ATTTTCAACAATATCTCAGAAGCAAAATTTGGTACGAGAACTCTCTACTGTGTAATACATC 360
Db 105 ValValAlaAsnLeuGlnArgGlnValSerSerLysIleValAlaSerGluTyrLeu 124
Qy 361 TGACCCCAACAGATATTTTGTTCATGTTATGAAAGGTTATGAAATCTCCAGAAATAGCT 420
Db 125 GluSerAsnLysAspLeuLeuAspIleLeuIleLeuGlyTyrGluAsnMetAspIleAla 144
Qy 421 CTAAATTTGGAGTAATATGTTAAGAGAAATGTCATCAGACCATCAACCATTTGCAAAATCAT 480
Db 145 LeuHisTyrGlyAlaMetLeuArgGluCysIleArgHisGlnSerIleAlaArgTyrVal 164
Qy 481 TTGTGTCGGAACAGAGTTTATGATTTTTCAGATATGTCGAAATGTCAACATTTGACATA 540
```

Db 165 LeuGluSerGluHisMetLysLysPhePheAspTyrIleGlnLeuProAsnPheAspIle 184
Qy 541 GCTTCAGATGCAATTCGCCACATTCAGGATTACTTACAGACATAAATTCGTCAGTGA 600
Db 185 AlaSerAspAlaSerAlaThrPheLysGluLeuThrArgHisLysAlaThrValAla 204
Qy 601 GAATTTTGGACAGCATATGATAGATATTTTTCAGTGAATAT--GAGAGTTTACTTCAT 657
Db 205 GluPheLeuSerAsnAsnTyrAspTTPhePheGluGluPheAsnSerArgLeuLeuSer 224
Qy 658 TCAGAAAATATGTGCAAAAAGACAGTCACTCAAGCTTCTCGGTGAACCTACTACTAGAT 717
Db 225 SerThrAsnTyrIleThrLysArgGlnAlaIleLysLeuLeuGlyAspMetLeuLeuAsp 244
Qy 718 AGACACAACTTCAATATGACAAATATACATCAGTAAACCTCAGAACCTCAATTAATG 777
Db 245 ArgSerAsnAlaAlaValMetMetArgTyrValSerSerLysAspAsnLeuMetIleLeu 264
Qy 778 ATGAACTGTCGAGACAAAGTCCGCAACATCCAGTTTCAGGCTTTCACGTTTTTAAG 837
Db 265 MetAsnLeuLeuArgAspSerSerLysAsnIleGlnIleGluAlaPheHisValPheLys 284
Qy 838 GTGTTTCTAGCCATCCTAAACAGCAGCCATCCTAGACATCCTCTCAAGAACCCAG 897
Db 285 LeuPheAlaAlaAsnLysAsnLysProGluValValAsnIleLeuValThrAsnArg 304
Qy 898 GCCAACTCATAGAGTTTCTCAGCAAGTTTTCAGAACGACAGGACCGAT 945
Db 305 SerLysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 320

RESULT 10
US-10-437-963-161075
; Sequence 161075, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161075
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET4530_60293C.1.pap
US-10-437-963-161075

Alignment Scores:
Pred. No.: 3,766-64 Length: 336
Score: 747.00 Matches: 146
Percent Similarity: 67.30% Conservative: 66
Best Local Similarity: 46.35% Mismatches: 99
Query Match: 41.43% Indels: 4
DB: 16 Gaps: 2
US-10-089-688-1 (1-1053) x US-10-437-963-161075 (1-336)
Qy 13 TTGGAAGTCTCAAAATCTCCAGACATTTGTAAGAATCTGAAGGAGCATGCT 72
Db 5 PheLysSerLysProAlaArgValValArgGlnThrArgGluLeuLeu 24
Qy 73 GTTCTGGAAAAGCAAGATTTCT-----GATAAAAAGCAGAAAAGGCTACAGAA 123

Db 25 PheLeuAspLeuHisSerGlySerArgGlyGlyAspAlaLysArgGluGluLysMetAla 44
Qy 124 GAAGTTTCCAAAATCTGGTTCCTCCATGAAGAAAATTCGTATGGCACAATAAGAAAAGAG 183
Db 45 GluLeuSerLysAsnIleArgGluLeuLysSerIleLeuTyrGlyAsnGlyGluSerGlu 64
Qy 184 CCTCAGACAGACAGTACGCTCAATTCGCTCAAGAACTCTATAATAGTGGGCTCCTTACG 243
Db 65 ProValThrGluAlaCysValGlnLeuThrGlnGluPhePheArgGluAsnThrLeuArg 84
Qy 244 ACCCTGGTACGTGATTTACAGCTCATTGACTTTGAGGCGCAAAAAGACGCTGGCTCAAAAT 303
Db 85 LeuLeuIleCysLeuProLysLeuAsnLeuGluThrArgLysAspAlaThrGlnVal 104
Qy 304 TTCACAAATATCTCAGAACAGACAAAATTTGTCAGAACCTCTACTGTTTGAATACATCTGC 363
Db 105 ValAlaAsnLeuGlnArgGlnValSerSerLysIleValAlaSerGluThrLeuGlu 124
Qy 364 ACCCAACAGAAATATTTTGTTCATGTTTGAAGGTATGAATCTCCGAAATAGCTCTA 423
Db 125 AlaAsnLysAspLeuLeuAspThrLeuIleSerGlyTyrGluAsnMetAspIleAlaLeu 144
Qy 424 AATTGGAATATGTTAGAGAAATGCATCAGACATGAACCACTTCGCAAAAATCATTTTG 483
Db 145 HisTyrGlySerMetLeuArgGluCysIleArgHisGlnSerIleAlaArgTyrValLeu 164
Qy 484 TGTGTCGGAACAGTTTTTATGATTTCTTCAGATATGCGAAATGTCAACATTTGACATAGCT 543
Db 165 GluSerAspHisMetLysLysPhePheAspTyrIleGlnLeuProAsnPheAspIleAla 184
Qy 544 TCAGATGCAATTTGCCACATTTCAAGGATTTACTTACAGACATAAATTCGTCAGTGCAGAA 603
Db 185 SerAspAlaSerAlaThrPheLysGluLeuLeuThrArgHisLysAlaThrValAlaGlu 204
Qy 604 TTTTGGACAGCATTTATGATAGATTTTTCAGTGAATATAG--AAGTTTACTTTCATCA 660
Db 205 PheLeuSerLysAsnTyrAspTTPhePheSerGluPheAsnThrArgLeuLeuSerSer 224
Qy 661 GAAAATTTATGTGCAAAAAGACAGACAGTCACTCAAGCTTCTCGGTGAACCTACTACTAGATAGA 720
Db 225 ThrAsnTyrIleThrLysArgGlnAlaIleLysPheLeuGlyAspMetLeuLeuAspArg 244
Qy 721 CACAACTTCAATATGACAAATATCATCATGATAAACCCTGAGAACCTCAAAATTAATCATG 780
Db 245 SerAsnSerThrValMetMetArgTyrValSerSerLysAspAsnLeuMetIleLeuMet 264
Qy 781 AACTGCTGCGAGACAAAAGTCCGACATCCAGTTTGAGGCTTTTCACGTTTTTAAAGTG 840
Db 265 AsnLeuLeuArgAspSerSerLysAsnIleGlnIleGluAlaPheHisValPheLysLeu 284
Qy 841 TTTGTAGCCCAATCTCAACAGACGACGCCATCTCTAGACATCTCTCTCAAGAACCCAGGCC 900
Db 285 PheAlaAlaAsnLysAsnLysProThrGluValValAsnIleLeuValThrAsnArgSer 304
Qy 901 AAACATCATAGATTTCTCAGCAAGTTTTCAGAACGACAGGACGGAT 945
Db 305 LysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 319

RESULT 11
US-10-425-114-54669
; Sequence 54669, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B

Db 121 Glu

QY 331 GGTACGAGAACTCCTACTGTTGAATACATCTGCACCCACAGAAATATTTGTTTCATGTTA 390
Db |||:||||| |||:||||| :|||||
110 AspGluAlaTyrCysValGlnTyrIleGluAsnHisPheAspLeuLeuAspPheLeu 129
QY 391 TTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGGAATAATGTTAAGAGAATGC 450
Db |||:||||| |||:||||| |||:||||| |||:|||||
130 ValValCysTyrIleAsnLeuGluValAlaLeuAsnCysGlyAsnMetLeuArgGluCys 149
QY 451 ATCAGACATGAACACATCTGCAAAATCATTTTGTGTCGGAACAGTTTATGATTTCTTC 510
Db |||:||||| |||:||||| |||:||||| |||:|||||
150 IleLysTyrProThrLeuAlaLysTyrIleLeuGluSerSerPheGluLeuPhePhe 169
QY 511 AGATATGCGAATGTCACATTTGACATAGCTTCAGATTCATTTGCCACATTCACAGGAT 570
Db |||:||||| |||:||||| |||:||||| |||:|||||
170 GlnTyrValGluLeuSerAsnPheAspIleAlaSerAspAlaLeuAsnThrPheLysAsp 189
QY 571 TTACTTACACACATTAATTCCTCAGTCGAGAATTTTGGACAGCAATATGATAGATTT 630
Db |||:||||| |||:||||| |||:||||| |||:|||||
190 LeuLeuThrLysHisGluAlaAlaValSerGluPheLeuCysSerHisTyrGluGlnPhe 209
QY 631 TTCAGTGAATATGAGAAGTTACTTCAATTCAGAAAATTTATGTGCAAAAAGACAGTCACTG 690
Db |||:||||| |||:||||| |||:||||| |||:|||||
210 PheGluLeuTyrThrArgLeuLeuThrSerThrAsnTyrValThrArgGlnSerVal 229
QY 691 AAGCTTCTCGGTGAACTACTACTAGATAGACAACTTCACAAATTTATGACAAAAGACAGTCACTG 750
Db |||:||||| |||:||||| |||:||||| |||:|||||
230 LysPheLeuSerGluPheLeuLeuGluAlaProAsnAlaGlnIleMetLysArgTyrIle 249
QY 751 AGTAACTGAGAACCTCAATTAATGATGAACCTGCTCGAGACAAAAGTCGCAACATC 810
Db |||:||||| |||:||||| |||:||||| |||:|||||
250 ValGluValSerTyrLeuAsnIleMetIleGlyLeuLeuLysPheSer**LysAsnIle 269
QY 811 CAGTTTGAGGCTTTCAAGTGTGTTTAAAGTGTGTTGAGCCAAATCCTAAACAGACGACGCC 870
Db |||:||||| |||:||||| |||:||||| |||:|||||
270 ArgIleCysAlaPheHisIlePheLysValPheValAlaAsnProAsnLysProArgAsp 289
QY 871 ATCTAGACATCTCTCTCAAGAACCGAGCCAACTCATAGATGTTCTTCAGCAAGTTTCAG 930
Db |||:||||| |||:||||| |||:||||| |||:|||||
290 IleIleGlnValLeuValAspAsnHisArgGluLeuLysLeuLeu----- 305
QY 931 AACGACAGGCGATTGTATGACGAGTTCCTGACGACGAGCAATTCCTCGGTCGATTTA 990
Db |||:||||| |||:||||| |||:||||| |||:|||||
306 -----GlyAsnLeuProThrSerLysGlyGluAspGluGln 317
QY 991 CGCGTTAAACCGCGACGCGTGGGATCAGGGATTTGAAGAGA 1032
Db |||:||||| |||:||||| |||:||||| |||:|||||
318 LeuGluGluGluArgAspLeuIleIleLysGluIleGluLys 331

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1614.5	89.5		341	2	I57997	hypothetical calci
2	1054.5	58.5		377	2	T16651	hypothetical prote
3	1006	55.8		338	2	T27129	hypothetical prote
4	766	42.5		329	2	T50117	mo25 homolog [limp
5	666.5	37.0		305	2	G71441	hypothetical prote
6	614	34.1		348	2	B84448	hypothetical prote
7	501	27.8		399	2	S34681	hypothetical prote
8	137.5	7.6		339	2	T33477	hypothetical prote
9	122	6.8		602	2	T3057	hypothetical prote
10	117	6.5		365	2	D97848	hypothetical prote
11	115	6.4		2819	2	A90551	conserved hypothet
12	113.5	6.3		1939	2	T18372	repeat organellar
13	113	6.3		2166	2	G70163	hypothetical prote
14	112.5	6.2		677	2	H64574	DNA topoisomerase

Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuTyrglyThrAsnGluLys 60
QY 181 GAGCCTCAGACAGAGCAGTAGCTCAACTTGGCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuValSerGlyLeuLeu 80
QY 241 AGACCCCTGGTAGCTGATTTACAGCTCATTTGAGGGCAAAAGACGCTGGCTCAA 300
Db 81 GlyThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuValAlaGln 100
QY 301 ATTTTCAACAATATCTCAGAGACAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
Db 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrlle 120
QY 361 TGCACCAACAGATATTTTGTTCATCTTATTTGAAGGGTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrgluSerProGluIleAla 140
QY 421 CTAATTTGTGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAGAAATCATT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIlelle 160
QY 481 TTGTGGTCGGAACAGTTTTATGATTTCTCAGATATGTCGAATGTCGAATTCACATTCACAT 540
Db 161 LeuTrpSerGluGlnPheTyrgluSerPheArgTyrgluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCTATTTGCCACATTCACAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAATTTTGGCAACAGCATTTATGATGATTTTTCAGTGAATATGAGAATGTTACTTCATTC 660
Db 201 GluPheLeuGluGlnHisTyrgluSerPheSerGluTyrgluLysLeuLeuHisSer 220
QY 661 GAAATATGATGACAAAGACAGTCACTGAGCTTCTCGGTGAACCTACTACTAGATAGA 720
Db 221 GluAsnTyrgluThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 240
QY 721 CACAACTTCAATATGACAAATATACATCAGTAACTGAGACCTCAAGATTAATCATG 780
Db 241 HisAsnPheThrIleMetThrLysTyrgluSerLysProGluAsnLeuLysLeuMet 260
QY 781 AACCTGCTGCAGACAAAGTCGCAACATCAGTTTGGAGGCTTTCACGTTTTTAAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCAATCTTAAACAGACGAGCCCATCTAGACATCTCTCAAGAACAGGCC 900
Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
QY 901 AAACCTCAGATGCTCCTCAGCAAGTTTCAGAACGACGACGAGGATTTGTATGAGCAGTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
QY 961 GTACCGACGACGAATTCCTCGGGTTCGATTTACGCGTTAAACCGCGACGCTGGGATCAGG 1020
Db 316 AspGluGlnPheAsnAspGluLysThrTyrgluValLysGln-----IleArg 331
QY 1021 GATTGTAGAGACACGCTCAGCAAGAGCT 1050
Db 332 AsnLeuLysArgAlaAlaGlnGlnGluAla 341

RESULT 2

Tl6651

Hypothetical protein R02E12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: Tl6651

R:Leimbach, D.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid R02E12.

A:Reference number: Z18554

A:Accession: Tl6651

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-377 <LEI>

A:Cross-references: UNIPROT:Q21643; EMBL:U53337; NID:g1255833; PID:g1255838; PIDN:AAA961

A:Experimental source: strain Bristol N2; clone R02E12

C:Genetics:

A:Gene: CESP:R02E12.2

A:Map position: X

A:Introns: 37/3; 146/2; 225/1; 315/3

C:Superfamily: Saccharomyces hypothetical protein YKL189w

Alignment Scores:	1.6e-73	Length:	377
Pred. No.:	1054.50	Matches:	205
Score:	77.44%	Conservative:	49
Percent Similarity:	62.50%	Mismatches:	61
Best Local Similarity:	58.49%	Indels:	13
Query Match:	2	Gaps:	2

US-10-089-688-1 (1-1053) x Tl6651 (1-377)

QY 1 ATCCCGTTCCTCGGTTTCGGAAGTCTCAAAATCTCCAGACAGACATTTGTGAAGATCTGAAG 60
Db 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20
QY 61 GAGACATGCTGCTTCTCGAAAAG-----CAAGAC 90
Db 21 GluValLeuThrIleLeuAspLysLeuProProLysLeuAspLysAspGlyAsnIle 40
QY 91 ATTCTGATATAAAGCAGAAAGGCTACAGAAAGTTCACAAATCTCCAAAATCTGGTGGCCATG 150
Db 41 GlnSerAspLysLysTyrgluLysAlaLeuAspLysValSerLysAsnValAlaMetIle 60
QY 151 AAAGAAATCTGATGCGCAAAATGAAAGAGAGCTCCAGACAGAA-----GCAGTA 201
Db 61 LysSerPheIleTyrgluAsnAspSerAlaGluProSerSerGluHisValValGlnVal 80
QY 202 GCTCAACTGCTCAAGAACTCTATATAGTGGGCTCTTAGCACCTCTAGTAGCTGATTTA 261
Db 81 AlaGlnLeuAlaGlnGluValTyrgluAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100
QY 262 CAGCTCATTCACCTTGGGCGCAAAAGAGCTCGCTCAAAATTTTCAACAATATTTCTCAGA 321
Db 101 ProLysPheGluPheGluCysLysLysAspValGlyGlnIlePheAsnAsnLeuLeuArg 120
QY 322 AGACAAATTCGTACGAACTCTCTACTGTTGAATATACATCTCGACCCACCAAGATATTTG 381
Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyrgluAlaArgProGluIleLeu 140
QY 382 TTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGAATAATGTTA 441
Db 141 IleGlnLeuValGlnGlyTyrgluProAspIleAlaLeuThrCysGlyLeuMetLeu 160
QY 442 AGAAGATGCTCAGACATGAACCACTTGCAGAAATCATTTTGGTGGGACAGCTTTTAT 501
Db 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTyrgluSerAspValPheTyrglu 180
QY 502 GATTCTTCAGATATGTCGAAATGTCACAAATGTCACATGTTGACATGCTCAGATGCTCCACA 561
Db 181 ThrPhePheLeuTyrgluValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200
QY 562 TTCAGAGATTACTTCAAGACATAAATTCGCTCAGTCCAGAAATTTTGGACACAGCATAT 621
Db 201 PheLysGluLeuThrThrArgHisLysAlaIleAlaGluPheLeuAspSerAsnTyrglu 220
QY 622 GATAGATTTTTCAGTGAATGAGAGCTTACTTTCATTTCAGAAATATTTGTGACAAAAAGA 681
Db 221 AspThrPhePheAlaGlnTyrgluAsnLeuLeuAsnSerLysAsnTyrgluValThrArgArg 240
QY 682 CAGTCACTGAGCTTCTCGGTGAACCTACTACTAGATAGACACAACTTCACAAATTTATGACA 741
Db 241 GlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArgHisAsnPheAsnThrMetThr 260

QY 742 AATACATCATGTAACCTGAGAACCTCAAAATTAATGTAACCTGCTCGAGACAAAGT 801
|||||
Db 261 LysTyrIleSerAsnProAspAsnLeuArgLeuMetGluLeuLeuArgAspLysSer 280
|||||
QY 802 CGCAACATCCAGTTGAGGCTTTTCACGTTTAAAGTGTGTTAGGCAATCCCTAAACAAG 861
|||||
Db 281 ArgAsnIleGlnTyrGluAlaPheHisValPheLysValPheValAlaAsnProAsnLys 300
|||||
QY 862 ACGAGGCCATCTAGACATCTCTCTCAAGAACAGGCCAAACTCATAGAGTTCTCTCAGC 921
|||||
Db 301 ProLysProIleSerAspIleLeuAsnArgAsnArgGluLysLeuValGluPheLeuSer 320
|||||
QY 922 AAGTTTCAGAACGACGAGCGGAT 945
|||||
Db 321 GluPheHisAsnAspArgThrAsp 328
|||||
RESULT 3
T27129
hypothetical protein Y53C12A.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27129
R:Kershaw, J.; Lennard, N.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z20315
A:Accession: T27129
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-338 <WIL>
A:Cross-references: UNIPROT:O18211; EMBL:Z99277; PIDN:CAB16486.1; GSPDB:GN00020; CESP:Y5
A:Experimental source: clone Y53C12A
C:Genetics:
A:Gene: CESP:Y53C12A.4
A:Map position: 2
A:Introns: 29/3; 103/3; 136/2; 215/1; 282/3
C:Superfamily: Saccharomyces hypothetical protein YKL189W
Alignment Scores:
Pred. No.: 8.67e-70 Length: 338
Score: 1006.00 Matches: 188
Percent Similarity: 77.96% Conservative: 56
Best Local Similarity: 60.06% Mismatches: 67
Query Match: 55.80% Indels: 2
Gaps: 1
US-10-089-688-1 (1-1053) x T27129 (1-338)
QY 13 TTTGGGAAGTCTCACAATCTCCAGCAGACATCTGTGAAGATCTGAGGAGCATCGCT 72
|||||
Db 6 PheGlyLysAlaAspLysThrProAlaAspValValLysAsnLeuArgAspAlaLeuLeu 25
|||||
QY 73 GTTCTGGAAGCAA-----GACATTTCTGATAAAAGCAGAAAAGGCTACAGAGAA 126
|||||
Db 26 ValIleAspArgHisGlyThrAsnThrSerGluArgLysValGluLysAlaIleGluGlu 45
|||||
QY 127 GTTTCGAAAATCTGGTGGCATGGAAGAAATCTGTATGGCACAATGAAAAGAGCCT 186
|||||
Db 46 ThrAlaLysMetLeuAlaLeuAlaLysThrPheIleTyrGlySerAspAlaAsnGluPro 65
|||||
QY 187 CAGACAGACAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCTTAGCACC 246
|||||
Db 66 AsnAsnGluGlnValThrGlnLeuAlaGlnGluValTyrAsnAlaAsnValLeuProMet 85
|||||
QY 247 CTGGTAGCTGATTACAGCTCATTCAGCTTTGAGGCGCAAAAGACGCTGCGTCAAAATTTTC 306
|||||
Db 86 LeuIleLysHisLeuHisLysPheGluPheGluCysLysLysAspValAlaSerValPhe 105
|||||
QY 307 AACAAATTTCTAGAGACAAATTTGGTACGAGAACTCTCTACTGTGTAATCATCTGCACC 366
|||||
Db 106 AsnAsnLeuLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrLeuAla 125
|||||
QY 367 CAACAGAAATTTTGTTCATGTTATTCGAAAGGCTATGAATCTCCAGAAATAGCTCTAAT 426
|||||

Db 126 ArgProGluLeuLeuLeuThrLeuLeuLeuGlyTyrGluGlnProAspIleAlaLeuThr 145
QY 427 TGTGGAATAATGTTAAGAGAAATGCATCAGACATCAACCACTTGCAGAAATCATTTTGTGG 486
|||||
Db 146 CysGlySerMetLeuArgGluAlaValArgHisGluHisLeuAlaArgIleValLeuThr 165
|||||
QY 487 TCCGAACAGTTTATGATTTCTTCAGATATGTCGAATATGTCGAATTTGACATTTGACATAGCTTCA 546
|||||
Db 166 SerGluTyrPheGlnArgPheValPheValGlnSerAspValPheAspIleAlaThr 185
|||||
QY 547 GATGCATTTCCCACTTCAAGGATTTACTTTACAGACATAAATTTGCTCAGTGCAGAAATTT 606
|||||
Db 186 AspAlaPheSerThrPheLysAspLeuMetThrLysHisLysAsnMetCysAlaGluTyr 205
|||||
QY 607 TTGGAACAGCATTTATGATATTTTTCAGTGAATATGAGAAGTTTACTTCAATTCAGAAAAAT 666
|||||
Db 206 LeuAspAsnAsnTyrAspArgPheGlyGlnTyrSerAlaLeuThrAsnSerGluAsn 225
|||||
QY 667 TATGTGACAAAAGACAGTCACCTGAAAGCTTCTCGGTGAACCTACTACTAGATAGACACAAC 726
|||||
Db 226 TyrValThrArgArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArgHisAsn 245
|||||
QY 727 TTCACAATTTATGACAAAATACATCATGATAACCTGAGAACCTCAAAATTAATGATGAACCTG 786
|||||
Db 246 PheSerThrMetAsnLysTyrIleThrSerProGluAsnLeuLysThrValMetGluLeu 265
|||||
QY 787 CTCGAGACAAAAGTCGCAACATCCAGTTTGTAGGCTTTTCACGTTTTTAAAGTGTGTTGTA 846
|||||
Db 266 LeuArgAspLysArgArgAsnIleGlnTyrGluAlaPheHisValPheLysIlePheVal 285
|||||
QY 847 GCCAATCTTAAAGACGCGCCCTTAGACATCTCTCCTCAAGAACGAGGCAAACTC 906
|||||
Db 286 AlaAsnProAsnLysProArgProIleThrAspIleLeuThrArgAsnArgAspLysLeu 305
|||||
QY 907 ATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGACGACGAT 945
|||||
Db 306 ValGluPheLeuThrAlaPheHisAsnAspArgThrAsn 318
|||||
RESULT 4
T50117
mo25 homolog [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50117
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25039
A:Accession: T50117
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-329 <SEE>
A:Cross-references: UNIPROT:Q9P7Q8; EMBL:AL157734; PIDN:CAB75774.1; GSPDB:GN00066; SPDB:1
A:Experimental source: strain 972h(-); cosmid c1834
C:Genetics:
A:Gene: SPDB:SPAC1834.06c
A:Map position: 1
A:Introns: 34/3; 185/3
C:Superfamily: Saccharomyces hypothetical protein YKL189W
Alignment Scores:
Pred. No.: 2.72e-51 Length: 329
Score: 766.00 Matches: 159
Percent Similarity: 68.37% Conservative: 55
Best Local Similarity: 50.80% Mismatches: 95
Query Match: 42.48% Indels: 4
Gaps: 2
US-10-089-688-1 (1-1053) x T50117 (1-329)
QY 1 ATGCCGTTCCCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGTAAGAAATCTGAAG 60
|||||
Db 1 MetSerPheLeuPheAsnLysArgProLysSerThrGlnAspValValArgCysLeuCys 20
|||||

```
QY 61 GAGAGCATGGCTCTTCTGGAAAGCAGACATTTCTGATAAAAGCAGAAAGCGCTACA 120
Db : : : : :
21 AspAsnLeuProLysLeuGluLeuAen-----AsnAspLysLys-----LysSerPhe 36
QY 121 GAGAGATTTCCAAAATCTGGTTGCCATGAAAGAAATTTCTGATGGCACAATGAAAAA 180
Db : : : : :
37 GluGluValSerLysCysLeuGlnAenLeuValSerLeuGlyThrAlaGluVal 56
QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db : : : : :
57 GluProAspAlaAspLeuValSerAspLeuSerPheGlnLeuTyrGlnSerAsnLeuPro 76
QY 241 AGCACCTGTGTAGTGAATTTAGCTCATTGATTCAGGGCAAAAAGACGCTGCTCAA 300
Db : : : : :
77 PheLeuLeuValArgTyrLeuProLysLeuGluPheGluSerLysLysAspThrGlyLeu 96
QY 301 ATTTTCAACAATATCTCAAGACAAATGTTGTTACGAGAACTCTACTGTTGAATACATC 360
Db : : : : :
97 IlePheSerAlaLeuLeuArgHisValAlaSerArgTyrProThrValAspTyrMet 116
QY 361 TGCACCCCAACAGAAATATTTGTTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCT 420
Db : : : : :
117 LeuAlaHisProGlnIlePheProValLeuValSerTyrTyrArgTyrGlnGluValAla 136
QY 421 CTRAAATGTGGATATGTTAAGAAATGTCATCAGACATGCAACCTTGCAGAAATCATT 480
Db : : : : :
137 PheThrAlaGlySerIleLeuArgGluCysSerArgHisGluAlaLeuAenGluValLeu 156
QY 481 TTGTGGTGGCAACAGTTTATGTTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db : : : : :
157 LeuAsnSerArgAspPheThrPhePheSerLeuIleGlnAlaSerSerPheAspMet 176
QY 541 GCTTCAGATGCTATGTCACATTCAGGATTTACTTACAGACATAAATTTGCTCAGTGCA 600
Db : : : : :
177 AlaSerAspAlaPheSerThrPheLysSerIleLeuLeuAenHisLysSerGlnValAla 196
QY 601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATGAGAGTTACTTCAATCA 660
Db : : : : :
197 GluPheIleSerTyrHisPheAspGluPhePheLysGlnTyrThrValLeuLeuLysSer 216
QY 661 GAAATATATGACAAAAGACAGTCACTGAAGCTTCTCGGTGAACACTACTACTAGATAGA 720
Db : : : : :
217 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGlyGluIleLeuAenArg 236
QY 721 CACAACCTTCAACATATGACAAAATACATCAGTAAACCTGAGAACCTCAATTAATGATG 780
Db : : : : :
237 AlaAsnArgSerValMetThrArgTyrIleSerSerAlaGluAenLeuLysLeuMetMet 256
QY 781 AACCTGCTGGCAGACAAAAGTCGCAACATCCAGTTGAGGCTTTTCAGTTTAAAGTG 840
Db : : : : :
257 IleLeuLeuArgAspLysSerLysAsnIleGlnPheGluAlaPheHisValPheLysLeu 276
QY 841 TTTGTAGCAATCTCTAACAAGCAGCAGCCCATCTAGACATCCCTCTCAAGAACCCAGGCC 900
Db : : : : :
277 PheValAlaAenProGluLysSerGluGluValIleGluIleLeuArgArgAsnLysSer 296
QY 901 AAATCATAGAGTTCTCTCAGCAAGTTTCAGAACGACGAGG 939
Db : : : : :
297 LysLeuIleSerTyrLeuSerAlaPheHisThrAspArg 309

RESULT 5
G71441
Hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Accession: G71441
R:Sevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
```

```
C.: Chalwatzis, N.
A>Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thali
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: G71441
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-305 <BEV>
A:Cross-references: GB:Z97343; NID:g2245073; PID:e327051; PID:g2245086
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Saccharomyces hypothetical protein YKL189w

Alignment Scores:
Pred. No.: 1-26e-43 Length: 305
Score: 666.50 Matches: 136
Percent Similarity: 65.68% Conservative: 63
Best Local Similarity: 44.88% Mismatches: 93
Query Match: 36.97% Indels: 11
DB: 2 Gaps: 2

US-10-089-688-1 (1-1053) x G71441 (1-305)
QY 124 GAAGTTTCCAAAATCTGTTGCCATGAAAGAAATTTCTGATGGCACAATGAAAAAGAG 183
Db : : : : :
8 GluLeuSerLysSerIleArgAspLeuLysLeuIleLeuTyrGlyAsnSerGluAlaGlu 27
QY 184 CCTCAGACAGACAGTAGCTCAACTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGC 243
Db : : : : :
28 ProValAlaGluAlaCysAlaGlnLeuThrGlnGluPhePheLysAlaAspThrLeuArg 47
QY 244 ACCCTGTGTAGCTGATTTTACAGCTCATTTGAGGGCAAAAAGACGCTGGCTCAAAAT 303
Db : : : : :
48 ArgLeuLeuThrSerLeuProAenLeuAenLeuGluAlaArgLysAspAlaThrGlnVal 67
QY 304 TTCACAAATATCTCAGACAGAAATGTTGTCAGAGAACTCTACTGTTGAATACATCTGC 363
Db : : : : :
68 ValAlaAenLeuGlnArgGlnGlnValAenSerArgLeuIleAlaAlaAspTyrLeuGlu 87
QY 364 ACCCAACAGCAATATTTGTTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTA 423
Db : : : : :
88 SerAsnIleAspLeuMetAspPheLeuValAspGlyPheGluAenThrAspMetAlaLeu 107
QY 424 AATTTGGAAATATGTTAAGAAATGTCATCAGACATGCAACCACTTGCAGAAAATCAATTTG 483
Db : : : : :
108 HisTyrGlyThrMetPheArgGluCysIleArgHisGlnIleValAlaLysTyrValLeu 127
QY 484 TGGTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCAACATTTGACATAGCT 543
Db : : : : :
128 AspSerGluHisValLysLysPhePheTyrTyrIleGlnLeuProAenPheAspIleAla 147
QY 544 TCAGATGCTATTGTCACATTTCAAGGATTTACTTACAGACATATAATTTGCTCAGTCAGAA 603
Db : : : : :
148 AlaAspAlaAlaThrPheLysGluLeuLeuThrArgHisLysSerThrValAlaGlu 167
QY 604 TTTTGGACAGCATATGATAGATTTTTCAGTGAATAT---GAGAAGTTACTTCACTCA 660
Db : : : : :
168 PheLeuIleLysAenGluAspTyrPhePheAlaAspTyrAenSerLysLeuLeuGluSer 187
QY 661 GAAATATGTCAGAAAAGACAGCTCACTGAAGCTTCTCGTGAACACTACTACTAGATAGA 720
Db : : : : :
188 ThrAsnTyrIleThrArgGlnAlaIleLysLeuGlyAspIleLeuLeuAspArg 207
QY 721 CACAACCTTCAATTTATGACAAAATATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
Db : : : : :
208 SerAsnSerAlaValMetThrLysTyrValSerSerMetAspAenLeuArgIleLeuMet 227
QY 781 AACCTGCTGGCAGACAAAAGTCGCAACATCCAGTTTGGGCTTTTCAGTTTAAAGTG 840
Db : : : : :
228 AsnLeuLeuArgGluSerSerLysThrIleGlnIleGluAlaPheHisValPheLysLeu 247
QY 841 TTTTGTAGCAATCTCTAACAAGCAGCAGCCCATCTAGACATCCCTCTCAAGAACCCAGGCC 900
Db : : : : :
248 PheValAlaAenGlnAsnLysProSerAspIleAlaAenIleLeuValAlaAenArgAen 267
```


A;Cross-references: EMBL:Z28189; NID:g486334; PIDN:CAA82032.1; PID:g486335; MIPS:YKL189w
A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:HYM1

A;Cross-references: SGD:S0001672

A;Map position: 11L

C;Superfamily: Saccharomyces hypothetical protein YKL189w

Alignment Scores:

Pred. No.:	7.3e-31	Length:	399
Score:	501.00	Matches:	110
Percent Similarity:	54.77%	Conservative:	68
Best Local Similarity:	33.85%	Mismatches:	137
Query Match:	27.79%	Indels:	10
DB:	2	Gaps:	3

US-10-089-688-1 (1-1053) x S34681 (1-399)

QY	1	ATGCCGTTCCCGTTGGAGTCTCAAAATCTCCAGAGACATTTGTGAAGATCTGAAG	60
DB	12	MetAlaPheTrpTrpLysAsnProLysThrProSerAspTyrAlaAArgLeuIle	31
QY	61	GAGAGCATGGCTGTCTGGAAGCAAGACATTTCTGATAAAAGAGCAAAAGCGTACA	120
DB	32	GluGlnLeuAsnLysPheSerSerProSerLeuThrGlnAspAsnLysAArgLysValGln	51
QY	121	GAAGAAGTTTCCAAAATCTGGTGGCCATGAAAGAAATCTGTATGCGACAAATGAAAA	180
DB	52	GluGluCysThrLysTyrLeuIleGlyThrLysHisPheIleValGlyAspThrAspPro	71
QY	181	GAGCTCAGACAGAGAGTGTCTGCTCAAGACATCTATATAGTGGGCTCCTT	240
DB	72	HisProThrProGluAlaIleAspGluLeuTyrThrAlaMetHisAArgAlaAspValPhe	91
QY	241	AGCACCTGTGAGTGAATTACAGCTCATTCAGCTTTGAGGCAAAAGAGCGTGCTCAA	300
DB	92	TyrGluLeuLeuLeuHisPheValAspLeuGluPheGluAlaAArgGluCysMetLeu	111
QY	301	ATTTTCAACATATCTCAGAGCAAAATGGTACGAGAACTCTCTACTGTGTATACATC	360
DB	112	IlePheSerIleCysLeuGlyTyrSerLysAspAsnLysPheValThrValAspTyrLeu	131
QY	361	TGCACCCAAAGATATTTTGTCTGATTTGTAAGAGGTATGAA-----	405
DB	132	ValSerGlnProLysThrIleSerLeuMetLeuArgThrAlaGluValAlaLeuGlnGln	151
QY	406	-----TCTCAGAAATAGCTCTAAATGTGGAATAATGTAAGAGATGATCAGACAT	459
DB	152	LysGlyCysGlnAspIlePheLeuThrValGlyAsnMetIleIleGluCysIleLysTyr	171
QY	460	GAACCACTTGCAGAAATCATTTTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTC	519
DB	172	GluGlnLeuCysAArgIleLeuLysAspProGlnLeuTyrLysPheGluPheAla	191
QY	520	GAATGTCAACATTTGACATAGCTTCAGATTTGCGACATTTGCCATTCAGAGTTTACTTACA	579
DB	192	LysLeuGlyAsnPheGluIleSerThrGluSerLeuGlnIleLeuSerAlaAlaPheThr	211
QY	580	AGACAT---AAATGCTCAGTCGCAATTTTGTGGAACAGCATTTATGAT-----AGATT	630
DB	212	AlaHisProLysLeuValSerLysLysGluPhePheSerAsnGluIleAlaAsnIleAArgPhe	231
QY	631	TTCAGTGAATGAGAACTTCTTCATTCAGAAATTTATGACAAAGAGACATGCTGCTG	690
DB	232	IleLysCysIleAsnLysLeuMetAlaHisGlySerTyrValThrLysAArgGlnSerThr	251
QY	691	AAGCTTCTCGGTGAATCTACTAGATAGACACATCTCACAATTTATGACAAAATACATC	750
DB	252	LysLeuLeuAlaSerLeuIleValIleArgSerAsnAsnAlaLeuMetAsnIleTyrIle	271
QY	751	AGTAACTGAGAACTCAATTAATGATGAACTGTCGCGAGACAAAGTCGCAATC	810
DB	272	AsnSerProGluAsnLysLeuIleMetThrLeuMetThrAspLysSerLysAsnLeu	291

QY	811	CAGTTTGGAGCCTTTTACGTTTGTAGCCAACTCTAACAGACGACGCC	870
DB	292	GlnLeuGluAlaPheAsnValPheLysValMetValAlaAsnProArgLysSerLysPro	311
QY	871	ATCTAGACATCTCTCAAGAACAGGCCAACTCATAGATTCCTCAGCAAGTTTCAG	930
DB	312	ValPheAspIleLeuValLysAsnArgAspLysLeuLeuThrTyrPheLysThrPheGly	331
QY	931	AACGACAGACGGAT	945
DB	332	LeuAspSerGlnAsp	336

RESULT 8

T33477

hypothetical protein T27C10.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T33477

R;Zhu, H.J.; Graves, T.; Hawkins, M.

submitted to the EMBL Data Library, October 1998

A;Description: The sequence of C. elegans cosmid T27C10.

A;Reference number: Z21354

A;Accession: T33477

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-339 <ZHU>

A;Cross-references: UNIPROT:Q9T2M2; EMBL:AF098504; PIDN:AAC67411.1; GSPDB:GN00019; CESP:1

A;Experimental source: strain Bristol N2; clone T27C10

C;Genetics:

A;Gene: CESP:T27C10.3

A;Map position: 1

A;Introns: 72/3; 120/3; 233/3; 295/1

Alignment Scores:

Pred. No.:	0.00748	Length:	339
Score:	137.50	Matches:	36
Percent Similarity:	43.89%	Conservative:	43
Best Local Similarity:	20.00%	Mismatches:	82
Query Match:	7.63%	Indels:	19
DB:	2	Gaps:	3

US-10-089-688-1 (1-1053) x T33477 (1-339)

QY	532	TTTGACATAGCTTCAGATGCAATTTGCCACATTCAGGATTTACTTACAAGACATAAATG	591
DB	109	PheAspValIleGlnGlyThrPheAspThrLeuGlnIleIlePhePheThrAsnHisGlu	128
QY	592	CTCAGTGCAGAAATTTTGGACACGATATGATAGATTTTTCAGTGAATATGAGAAGTTA	651
DB	129	SerAlaAsnAsnPheIleLysAsnAsnLeuProArgPheMetGlnThrLeuHisLysLeu	148
QY	652	CTTCATTCAGAAATTTATGTGCAAAAAGACAGCTCACTGAAGCTTCTCGGTGACTACTA	711
DB	149	IleAlaCysSerAsnPhePheIleGlnAlaLysSerPheLysPheLeuAsnGluLeuPhe	168
QY	712	CTAGATAGACACACTTCACAAATTTATGCAAAAATACATCAGTAAACCTGGAACCTCAAA	771
DB	169	ThrAlaGlnThrAsnTyrGluThrArgSerLeuTrpMetAlaGluProAlaPheIleLys	188
QY	772	TTATATGATGAACCTGTCGCGAGACAAAGTCGCAACATCCAGTTTGGAGCTTTTCAGTT	831
DB	189	LeuValValLeuAlaIleGlnSerAsnLysHisAlaValArgSerAArgAlaValSerIle	208
QY	832	TTTAAGTGTGTTAGCCAATCTTAAACAGCAGCGCCATCTCTAGACATCTCTCTCAAG	891
DB	209	LeuGluIlePheIleAArgAsnProArgAsnSerProGluValHisGluPheIleGlyArg	228
QY	892	AACAGCGCAACCTCATAGATTTCTCAGCAAGTTTCTCAGAACGACGACGAGCGATTGTATG	951
DB	229	AsnArgAsnValLeuIleAlaPhe-----ACGAGAAATTCGCGGTGATTTACGC	993

Db 238 PheAsnSerAlaProIleHisTyThrGlnGlySerProAsnGluLysGluAspAlaGln 257
Qy 994 GTTAAACCGCGAGCGCTGGATCAGG-----GATTGTAAGAGACAGCAGCTCAGCAGAA 1047
Db 258 TyAlaAlaArgMetAlaTyLysLeuLeuAsnTrpAspMetGlnArgProPheThrGlnGlu 277
RESULT 9
T03057
hypothetical protein 032R - Chilo iridescent virus
C:Species: Chilo iridescent virus
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03057
R:Bahr, U.; Tidona, C.A.; Barai, G.
Virus Genes 15, 235-245, 1997
A>Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101
A:Reference number: Z14834; MUID:98141693; PMID:9482599
A:Accession: T03057
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-602 <BAH>
A:Cross-references: UNIPROT:O55720; EMBL:AF003534; NID:G2738385; PIDN:AAB94431.1; PID:94
C:Superfamily: Chilo iridescent virus hypothetical protein 032R
Alignment Scores:
Pred. No.: 0.122 Length: 602
Score: 122.00 Matches: 75
Percent Similarity: 38.94% Conservative: 57
Best Local Similarity: 22.12% Mismatches: 145
Query Match: 6.77% Indels: 62
Gaps: 2
US-10-089-688-1 (1-1053) x T03057 (1-602)
Qy 10 CCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTCTGAAGAATCTGAAG----- 60
Db 107 ProPheLeuProAsnPheMetLysProTyAspLeuIleLysAsnValArgAlaAspPro 126
Qy 61 -----GAGAGCATGCTGTCTCGAAAGCAAGAC 90
Db 127 ArgGlnLysAsnProPheValThrLeuIleGluLysSerGlySerGluLysMetArg 146
Qy 91 ATTTCTGATATAAAGCAGAAAGGCTACA-----GAGAAGTT 129
Db 147 LeuThrThrLysLysSerLysProLysThrPheSerAspMetAlaLeuPheGluTyIle 166
Qy 130 TCCAAAAATCTGGTGGCCATGAAGAATCTGTATGGCCAAATGAAGAAGACCTCAG 189
Db 167 AsnSerAsnMetThrLeuAlaGluLeuIleSerIleAsnValLysGlyLysArgGlnThr 186
Qy 190 ACAGAAGCAGTAGCTCAACTTGTCTCAAGAATCTTATAATAGTGGG-----CTCCTTAGC 243
Db 187 ArgValThrLeuAsnGluGluIleGlnLysIleLysAlaSerAspPheLysIleIleAsn 206
Qy 244 ACCCTGTAGCTGATTACAGCTCATGTCTTGGGGCAAAAGACGTGGCTCAAAAT 303
Db 207 SerLeuLeuAsnGlnLeuMetIleAlaIleLeuIleAlaGlnLysLysLeuSerPheIle 226
Qy 304 -----TTCAACATATCTCAGAGACAAATTTGGTACGAGAACTCTCTACT 348
Db 227 HisAsnAspLeuHisPheAspAsnValLeuIleCysLysCysLeuGlnArgThr----Phe 245
Qy 349 GTTGAATACATCTGCACCAACAGAAATATTGTTTCATGTATTGTAAGAGGTATGAATCT 408
Db 246 MetLeuTyValPheGluTyThrAsnSerIleCysTyThrAlaLeuLeuProThrTyGlyTy 265
Qy 409 CCAGAATAGCTCTAAATTTGGTAATATGTTAAGAAATGCATCAGACATGAACCACTT 468
Db 266 TyrProIleIleAspTyThrPheSerPheSerGluAspLeuIleGlyGlyProLeu 285
Qy 469 GCAAAATCATTTTGGTGGCAGACAGTTTATGATTTCTTCAGATATGTCGAATGTCA 528
Db 286 LeuThrGlyIleHisAsnAsnLysGlyTyThrMetAsnHisGlnTy----- 301

Qy 529 ACATTTCACATAGCTTCAGATGCATTTGCCACATTCACGATTTACTTACAGACATATAA 588
Db 302 -----AspGluPheThrAspPheLysThrMetLeuThrArgLeuSer 315
Qy 589 TTGCTCAGTCAGCAAAATTTTGGGAACAGCATTATTATGATAGATTTTTCAGTGAATATGAGAAG 648
Db 316 TyrSerGlyTyThrGlnPheGlyLeuAspLysLysAspAlaPhe----- 329
Qy 649 TTACTTTCATTTCAGAAAAATTATGTGCAAAAAGACAGTCACTG-----AAGCTTCTC 699
Db 330 -----GlnSerLeuIlePheAspLysLeuIle 338
Qy 700 GGTGAACCTACTACTAGATAGACACAACTTCACAAATTATGACAAATATCAGTAACCT 759
Db 339 SerLysLeuProIleAspLysGlnThrGlyTrpAspGluThrLysAspIleSerValSer 358
Qy 760 GAGAACCTCAAAATTAATGATGAACCTGCTCGAGAC-----AAAAGTCGCAACATC 810
Db 359 LysGlnLeuValArgHisIleArgIlePheValAspAspTyThrLysSerIleAsnArg 378
Qy 811 CAGTTTGGAGCCCTTTCACCGTTTTTAAAGTGTGTTGTAGCCCAATCTTAACAAGACGACGCC 870
Db 379 Glu---SerPheGlnLysTyThrAspTyThrGluMetValAspMetIleGlySerLeuIle 397
Qy 871 ATCTAGACATCTCTCAGAACACGAGCCAAATCTCATAGATTCCTCAGCAAGTTT 927
Db 398 IleLeuProLeuArgLysLysAsnThrGluAsnLeuValGluThrLeuAlaIlePhe 416
RESULT 10
D97848
hypothetical protein RC1188 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: D97848
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A>Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: D97848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-365 <KUR>
A:Cross-references: UNIPROT:Q92GD5; GB:AE006914; PIDN:AAL03726.1; PID:g15620317; GSPDB:G
C:Genetics:
A:Gene: RC1188
Alignment Scores:
Pred. No.: 0.286 Length: 365
Score: 117.00 Matches: 74
Percent Similarity: 36.31% Conservative: 60
Best Local Similarity: 20.05% Mismatches: 127
Query Match: 6.49% Indels: 108
Gaps: 2
US-10-089-688-1 (1-1053) x D97848 (1-365)
Qy 46 GTCAAGAACTCTGAAGGAGAGAGATGGTGTCTCGAAAAGCAAGACATTTCTGATAAAAAA 105
Db 60 IleSerAsnLeuLysGluThrAlaGluThrLysLeuLysHisTyThrGluValGluGlnSer 79
Qy 106 GCAGAAAAGCTACAGAGAAGTTTCCAAAAATCTGGTTGCCATGAAGAATTTCTGTAT 165
Db 80 IleGluLysAlaArgValGlnGluGluPheLeuLysSerGlnArgAlaIleLysLeu 99
Qy 166 GCAGAAAATGAAGAGAGCTCAGACAGACGAGTACCTCAACTTGTCTCAAGAACTTAT 225
Db 100 AlaGluGluLysGlu-----IleAlaGlnArgIleAlaAlaLeuAsn 114
Qy 226 AAT-----AGTGGGCTCTTAGCACCCCTGGTAGCTGATTTCAGCTCATTGACTTTGAG 279
Db 115 AsnLeuHisAsnGluPheIleLysAsnIleThrLysAspThrLysArgIleGlu-----Glu 133


```
Qy 634 AGTGAATAGAGAGTTACTTTCATTCAGAAAATTATGTGACAAAAGACAGACTCACTGAG 693
Db 1667 AlaleuPheLysValLeuSerSerAsnLeuLeuSerThrGlnSer-----1682
Qy 694 CTTCTCGGTGAATCTACTAGATAGACACAACTTCACA-----ATTATG 738
Db 1683 -----AspGlnSerSerPheThrLysSerSerLeuLeu 1695
Qy 739 ACAAAATACATCAGTAACCTGAGAACCTCAAAATTAATGATGAACCTGTCGAGACAAA 798
Db 1696 ArgValPheLeuArgLysGluLysAlaLeuGluLeuMet-----LeuLysLysGln 1712
Qy 799 AGTCGCAACATCCAGTTT-----GAGGCGCTTTCACGTTTAAAGGTG 840
Db 1713 PheArgGlnPheGluLeuHisAspLysValAspAlaLeuAsnLeuLys---1731
Qy 841 TTTTAGCCCAATCTCAACAGACGAGCCCATCTAGACATCTCTCCCTCAAGAAC-----894
Db 1732 PheIlePheArgAspAsnGlnValIleAspPheIleGluThrIleIleAsnArgValVal 1751
Qy 895 -----CAGGCCAAACTCATAGATTCTCTC 918
Db 1752 SerGlnLysGluGluTyrSerLysLeuAsnSerTyrGlnGluIleLeuTyrLysPheLeu 1771
Qy 919 AGCAAGTTTCAGACGACGAGCGAGTGTATGAGCAGTTCCTGTA-----963
Db 1772 SerSerAsnLysAsnGluThrLeuAsnPhePheLysLysGlnIleGlyGluLeuLysAsn 1791
Qy 964 -----CCGACGACGAATTCGCGGTGCGATTACCGGTATAA 999
Db 1792 SerLysLeuLeuAsnProLeuAlaLysSerPheIleGluSerGluLeuLys 1808

RESULT 12
T18372
repeat organellar protein - Plasmodium chabaudi
C:Species: plasmodium chabaudi
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18372
R:Werner, E.B.; Taylor, W.R.; Holder, A.A.
Mol. Biochem. Parasitol. 94, 185-196, 1998
A:Title: A Plasmodium chabaudi protein contains a repetitive region with a predicted spe
A:Reference number: 218922; MUID:98418765; PMID:9747969
A:Accession: T18372
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1939 <WER>
A:Cross-references: UNIPROT:Q25662; EMBL:U43145; NID:g1151157; PID:g1151158; PIDN:AAC634

Alignment Scores:
Pred. No.: 0.593 Length: 1939
Score: 113.50 Matches: 57
Percent Similarity: 38.78% Conservative: 57
Best Local Similarity: 19.39% Mismatches: 113
Query Match: 6.30% Indels: 67
DB: 2 Gaps: 11

US-10-089-688-1 (1-1053) x T18372 (1-1939)
Qy 52 AATCTGAAGGAGAGAGTGGCTGTTCTGCAAAAGCAGACATTTCTGATAAAAGCAGAA 111
Db 1433 AsnLeuAsnLysAsnValGluAspLysThrGlyAspAspIleAsnCysGluLysAsnAsn 1452
Qy 112 AAGGCTACAGAGAAGTTTCC-----AAAAATCTGGTGGCCATGAAGAAGAAATTTCTGTAT 165
Db 1453 AspGlnAlaLysGluIleSerTyrLeuLysAspGluIleLysLysIleSerMetLeuTyr 1472
Qy 166 GGC-----ACAATGAAAGAGCCCTCAGACAGACAGCTAGCTCAACTTGTCTCAAGAA 219
Db 1473 GlyGluGluLeuAsnArgLysAsnSerTyrAspGluLysValLysAsnLeuThrAsnGlu 1492
Qy 220 CTCTATAATAGTGGCTCCTTAGCACCCCTGTAGTACCTGATTTACAGCTCATTTGACTTTGAG 279
Db 1493 Leu-----LysGluLeuLysIleArgAsnLysLys 1502
```

```
Qy 280 GGCACAAAAGACGCTGGCTCAAAAT-----303
Db 1503 GlyGluGluAlaIleAlaGluLeuAsnLysLeuLysAsnIleLysGluLysAsnLysSer 1522
Qy 304 -----TTCAACAATATTCTCAGACGACAAATTTGGTACGAGA 339
Db 1523 ValLysGlnAsnAspGluSerSerSerAsnAsnIleIleThrLysAsp---GlyAspLys 1541
Qy 340 ACTCTACTCTTGAATACATCTGCACCCCAACAGAAATATT-----TTGTTTC 384
Db 1542 ThrPro-----GluTyrValSerAsnAspAspLysIleGlnLysAspTyrLysAlaAsn 1559
Qy 385 ATGTTATTGAAAGGTATGAATCTCAGAAATAGCTCTAAATTTCTGGAATAATTTTAAAGA 444
Db 1560 LeuValLeuLysLysGluLysProAspLeuTyrAspAsnIleAsnSerLeuGluLys 1579
Qy 445 GAATGCATCAGACATGAACCATCTTGCAAAA-----ATCATTTTGTGGTGGACACAGTTT 498
Db 1580 GluAsnPheArgValMetSerIleValLysGluAsnLysAsnValGlnAsnAspLysIle 1599
Qy 499 TATGATTCTTCAGATATGCGAAATGTCACAAATTTGCACATTTGCAGATGCAATTTGCC 558
Db 1600 ValGlyIleTyrSerTyrPheLysLysCysGluLysGluLeuLysAsnAspMetLeuVal 1619
Qy 559 ACATTCAGGATTTACTTACAGACATAAATGCTGCTGAGTCAGAAATTTTGGACACAGAT 618
Db 1620 Ile-----CysLeuValLeuLysAspIleLeuSerIleLeuPheLeuAsnAspAsn 1636
Qy 619 TATGATAGATTTTTCAGTGAATATGAGAGCTTACTTCATTCAGAAAAATTTATGTCACAAA 678
Db 1637 PheValAsnLeuPheGluLysIleAspLysIleLeuTyrLysGlnMetTyrIleProThr 1656
Qy 679 -----AGACAGCTCACTG 690
Db 1657 GluIleArgIleLeuPheLeuArgTyrPheSerPheLeuAspLysLeuArgAsnTyrVal 1676
Qy 691 AAGTTCTTCGGTGAATCTACTACTAGAT-----AGACACAACTTCACAAATTTATG 738
Db 1677 LysCysValAsnGluGluTyrValAsnAsnGluArgTyrGluTyrSerTyrPalaLeuPhe 1696
Qy 739 ACAAAATACATCAGTAACCTGAGAACCTGAGAACCTCAAAATTAATGATG 780
Db 1697 GlnThrTyrLeuGluThrAlaSerAsnLeuLysGluMet 1710

RESULT 13
G70163
hypothetical protein BB0512 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: G70163
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: G70163
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2166 <KLE>
A:Cross-references: UNIPROT:O51465; GB:AB001153; GB:AB000783; NID:G2688419; PIDN:AAC6687
A:Experimental source: strain B31

Alignment Scores:
Pred. No.: 0.553 Length: 2166
Score: 113.00 Matches: 80
Percent Similarity: 43.30% Conservative: 72
Best Local Similarity: 22.79% Mismatches: 116
Query Match: 6.27% Indels: 84
DB: 2 Gaps: 21
```


398 PheLysIleLysSerGluSerPheLysLeuSerPheLysLeuLysGluLysGlyPhe 417
QY 556 GCCACATTCAGGATTACTTACAGACATATAAATGCTCAGTGCAGAAATTTTTCGAACAG 615
Db 418 LeuGluLeuGluLeuLeuLeuGlnGlyLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 437
QY 616 CATTATGATAGATTTC----- 633
Db 438 GluLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 457
QY 634 -----ACTGAATATGGAAGTACTT 654
Db 458 LysLysIleGluLysProSerProLysProLysGluSerAlaPheIleProLeuLeu 477
QY 655 CATTGAGAAATATGACAAAGACAGTCACTGAGCTTCCGGTGAACACTACTACTA 714
Db 478 GluSerGluGly-----IleGlyArgProSerThrTyrAlaSerPheLeuAspLeuLeu 496
QY 715 GATAGACACAACTTCACAATATTATGACAAA----- 744
Db 497 LysArgLysTyrIleSerIleAspThrLysThrAsnAlaIleThrProThrSerGlnGly 516
QY 745 -----TACATCAGTAAACCTGAGAACCTCAAAATTAATGATGAACTGCTG 789
Db 517 LeuGluValIleSerPhePheLysAspLysGluValAspPheIleAlaLeuThrSer 536
QY 790 CGAGACAAAAGTCCCAACATCCAGTTTGGAGCCCTTTCACGTTTTTAAGGTGTTGTAGCC 849
Db 537 LysAspLysSerLys-----LeuGly 543
QY 850 AATCCTAACAGACG---CAGCCCATCTAGACATCTCTCTCAGAACACCGCCAAA--- 903
Db 544 AsnThrThrLysGlnPheGluGluCysLeuAspLeuIleMetArgGlyGluAlaSerTyr 563
QY 904 -----CTCATAGAGTTCTCTCAGCAAGTTTCAGAACGAC-----AGGACG 942
Db 564 GluLysPheMetLeuGluValIleSerLysLeuLysSerThrAlaLysPheTyrGlnThr 583
QY 943 GATTGTATGAGCATTCGCTACCGACGACGAAATTCCTCGGTGCGATTTACGGTTAAACCG 1002
Db 584 GlnSerThrAspAsnMetProThrAspLys-----GlnLeuGluLeuIleAsp 600
QY 1003 CGGACGGTGGATCGAGGATTTCAGACGACCGCTCAGCAA 1044
Db 601 LysIleCysLysAspLysLysLeuGlnLysProSerGlnGlu 614

RESULT 15
T18489
hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18489
R/Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A/Reference number: Z18935
A/Accession: T18489
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4981 <LAW>
A/Cross-references: UNIPROT:O77372; EMBL:Z98551; NID:e1331903; PID:e1331910; PIDN:CAB111
C/Genetics:
A/Map position: 3
A/Note: C0820w

Alignment Scores:
Pred. No.: 0.984 Length: 4981
Score: 111.00 Matches: 59
Percent Similarity: 36.73% Conservative: 49
Best Local Similarity: 20.07% Mismatches: 104
Query Match: 6.16% Indels: 82
DB: 2 Gaps: 10

US-10-089-688-1 (1-1053) x T18489 (1-4981)

QY 298 CAAATTTTCAACATATTCTCAGACACAATTTGGTACGAGAACTCCTACTGTTGAATAC 357
Db 1521 LysIlePheLeuAsnTyrIleArgLysGluLysGluHisIleTyrAsnPheGlnThrTyr 1540
QY 358 ATCTGCACCCCAACAGAAATATTTTGTTCATGCTTTTAAAGGGTATGAATCT----- 408
Db 1541 IleThrThrPhePheThrLeuLeuGluMetLeuLysAsnIleGlnSerAspLysAsp 1560
QY 409 ---CCAGAAATAGCTCTAAATTTGTGAATAATGTTAAGAGAAATGCATCAGACATGAACCA 465
Db 1561 TyrGluLysThrLeuTyrAsnIleGlyIleIleAsnMetMetCysValAlaValLysPhe 1580
QY 466 CTTGCA-----AAATCATTT----- 480
Db 1581 LeuLysGluLysCysProAsnLysIleIleSerIlePhePheLysLeuLeuGluVal 1600
QY 481 -----TTGTGGTCCGGAACAG-----TTTATGATTTCTTCAGATAT 516
Db 1601 ProIleProLeuGlyGluGluGluAspLysValValThrLeuTyrLysAsnTyrSerTyr 1620
QY 517 GTCGAAATGTCAACATTTGACATAGCTTTCAGATGCAATTTGCCACATTCAGAGATTTACTT 576
Db 1621 IleTyrArgThrHisPhe-----PheAsnLeuTyrLysAspLysCys 1634
QY 577 ACAAGACATAAATGCTCAGTCAGAGAAATTTTGGAAACAGCATTATGATAGATTTTTCAGT 636
Db 1635 Thr-----ThrLeuTyrThrLysAsnCysMet 1643
QY 637 GAATATCAGAGTACTTCTCAGAAAATATGTGACAAAAGACAGCTCAGTGAAGCTT 696
Db 1644 GluTyrIleLysMetMetArgAsnAsnIleTrpLeuAsnLysLysLysAsn----- 1660
QY 697 CTCGGTGAATCTACTAGTATAGACACAACTTCACAAATATGACAAAATACATCAGT--- 753
Db 1661 ---AspSerGlnValAspPheAsnAsnPheIleIleLeuSerAsnMetTyrAsnAsn 1678
QY 754 -----AACTGAGAACCTCAAAATTAATGATGAACCTGCTGCGAGACAAAGT 801
Db 1679 SerAspThrPheLysLysLysAsnIleGlnAsnValLeuGlnTyrLeuAsnAspTyrGln 1698
QY 802 CGCAACATCCAGTTTGAGGCTTTCAGTTTAAAGGTGTTTGTAGCCAAT----- 852
Db 1699 GlnValTyrLysLysAsnValValHisIlePheAsnMetTyrIleSerAsnValThr 1718
QY 853 -----CCTAACAGACG 864
Db 1719 HisLysGluMetGluAsnGlyMetTyrArgLysGluGlyAsnLysArgGluAsnAsnThr 1738
QY 865 CAGCCCATCTAGACATCTCTCTCAAGAACAG----- 897
Db 1739 TyrAspLysIleAsnIleAspAspLysAsnAsnPheGlyGlnProIleTyrLysTyrVal 1758
QY 898 GCCAAACTCATAGAGTTCTCAGCAAGTTTCAGAACGACGAGGAGGATTTGATGACAGT 957
Db 1759 AspLysLeuThrSerLysArgAsnThrLeuAspAsnSerIleAsnCysLeuGluSer 1778
QY 958 TCCGTACCGACGACGAATTCCTCGGGTCGATTTCACGGTTAAA 999
Db 1779 SerTyrValThrLysGlySerAsnThrLysIleArgIleLys 1792

Search completed: November 10, 2004, 20:52:59
Job time : 74.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:33:32 ; Search time 236.5 Seconds
(without alignments)
5123.632 Million cell updates/sec

Title: US-10-089-688-1

Perfect score: 1803

Sequence: 1 atgcgcgtccgtttgggaa.....cagctcagcagaagcttaa 1053

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0	10.0	0.5
Xgapext 0.5	0.5	7.0
Delop 6.0	7.0	7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPRO.spool.p/US10089688/runat.10112004.152246.4843/app.query.fasta_1.1223
-DB=Uniprot.02 -QPMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089688 @CGN 1.1.244 @runat.10112004.152246.4843 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot.02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1635.5	90.7	341 1	CB39 HUMAN
2	1619.5	89.8	341 2	Q8VDZ8
3	1614.5	89.5	341 1	CB39 MOUSE
4	1614.5	89.5	341 2	Q6IPF2
5	1614.5	89.5	341 2	AAH72045
6	1590.5	88.2	341 2	Q724X0
7	1537.5	85.3	341 2	Q6IQI2
8	1537.5	85.3	341 2	AAH71393
9	1339.5	74.3	337 2	Q6PAB4
10	1339.5	74.3	337 2	AAH60384
11	1298.5	72.0	337 1	C39L MOUSE
12	1297.5	72.0	337 1	C39L HUMAN
13	1287.5	71.4	337 2	AAQ93064
14	1287	71.4	343 2	Q803V8
15	1163.5	64.5	332 2	Q7PF07
16	1163.5	64.5	335 2	Q7PRN5

17	1123	62.3	339	1	MO25 DROME
18	1123	62.3	339	2	AAH75031
19	1054.5	58.5	377	2	Q7Z2A5
20	1054.5	58.5	636	2	Q21643
21	1006	55.8	338	1	MO2M CABEL
22	944.5	52.4	205	2	Q8K3I2
23	766	42.5	329	1	YFV6 SCHPO
24	764.5	42.4	321	1	DE76 CHLPR
25	726.5	40.3	390	2	Q6CAN7
26	711.5	39.5	343	1	MO2N ARATH
27	711.5	39.5	343	2	AAH16128
28	705.5	39.1	343	1	MO2M ARATH
29	705.5	39.1	343	2	AAH06959
30	686	38.0	338	2	Q6BSQ5
31	669.5	37.1	345	2	Q8L9L9
32	651.5	36.1	337	2	Q7XIQ4
33	623	34.6	384	1	HYMA EMENI
34	614	34.1	348	1	MO2L ARATH
35	614	34.1	348	2	AAH24657
36	584	32.4	370	2	Q873K5
37	571	31.7	364	2	Q6CWP5
38	560	31.1	362	2	Q757C0
39	560	31.1	362	2	AAH52777
40	511	28.3	422	2	Q6FWG7
41	501	27.8	399	1	HYM1 YEAST
42	455.5	25.3	332	2	Q7QTG6
43	327	18.1	103	2	Q8K038
44	283.5	15.7	509	2	Q7YYL6
45	140.5	7.8	1111	2	Q9U0K5

ALIGNMENTS

RESULT 1

CB39_HUMAN ID CB39 HUMAN STANDARD; PRT; 341 AA.

AC Q9Y376; DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39 (Mo25 protein) (CGI-66).
DB Names=CAB39; Synonyms=Mo25;
GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Townestein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RA Jin W., Shi J., Ren S., Gu J., Fu S., Huang Q., Dong H., Yu Y., Fu G.,
RA Wang Y., Chen Z., Han Z.;
RT "A novel gene expressed in the human hypothalamus."
RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC TISSUE=Ductenium;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Townestein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 10-339.
RX PubMed=14730349; DOI=10.1038/nsm716;
RA Milburn C.C., Boudreau J., Deak M., Alessi D.R., van Aalten D.M.;
RT "Crystal structure of Mo25 alpha in complex with the C terminus of the
RT pseudo kinase STE20-related adaptor.";
RL Nat. Struct. Mol. Biol. 11:193-200(2004).
CC -!- FUNCTION: Together with the STE20-related adaptor-alpha (STRAD
CC alpha) pseudo-kinase, forms a regulatory complex capable of
CC stimulating the activity of Stk11.
CC -!- SIMILARITY: Belongs to the Mo25 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF151824; AAD34061.1; -;
DR EMBL; AF1513536; AAF14873.1; -;
DR EMBL; BC020570; AAH20570.1; -;
DR PDB; 1UPK; X-ray; A=1-341.
DR IntAct; Q9Y376; -;
DR GeneW; HGNC:20292; CAB39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
KW 3D-structure.
SQ SEQUENCE 341 AA; 39869 MW; EC710A528B6F9811 CRC64;

Alignment Scores:
Pred. No.: 6,78e-114 Length: 341
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 1 Gaps: 2
US-10-089-688-1 (1-1053) x CB39_HUMAN (1-341)
QY 1 ATGCGGTTCCCGTTGGGAGGCTCACAATCTCCAGCAGACATTTGTGAAGAATCTGAAG 60
DB 1 MetProPheProPheGlyysSerHisLysSerProAlaAspIleValIysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTCGAAAGCAAGACATTTCTGATATAAAGACGAGAAAGCTACA 120
DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAAGTTTCCAAAATCTGCTGCCATGAAGAAGAAATCTCTATGGCACAATCAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuLysGlyThrAsnGluLys 60
QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysAsnSerGlyLeuLeu 80
QY 241 AGACCCCTGGTAGCTGATTTACAGCTCATTTGACTTTGAGGCGGCAAAAAGACCTGCTCAA 300
DB 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100

RESULT 2

Q8VDZ8
ID Q8VDZ8 PRELIMINARY; PRT; 341 AA.
AC Q8VDZ8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cab39 protein.
GN Name=Cab39;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

QY 301 ATTTTCAACAATATTTCTCAGAGACAAATTTGTCAGAGACCTCTACTGTTGAATACATC 360
DB 111PheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCCAACAATATTTTGTTCATGTTATTAAGAGGTATGAATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAAATTTGGGAATATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAGAAATCAT 480
DB 141 LeuAsnGlyIleMetLeuArgGlnCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGTGTCGGAACAGTATTTATCATTTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
DB 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCAATTTGCCACATTTCAAGGATTTACTTACAGACATATAAATGCTCAGTGA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAATTTTTCGACAGCAGCATTTATGATATGATATTTTTCAGTGAATATGAGAAGTTACTTCA 660
DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
QY 661 GAAATATTTGTGACAAAAGACAGTCACTGAGCTTCTCGTGAACTACTACTAGATAGA 720
DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGlyLeuLeuLeuAspArg 240
QY 721 CACAATTTTCAATTTATGACAAATATCATCAGTAAACCTCAGAACCTCAAATTAATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACCTGCTCGGAGACAAAAGTCGCAACATCCAGTTTTCAGGCGCTTTTCACGTTTAAAGGTG 840
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCCAATCTTAACAGACGCGCCATCTAGACATCTCTCTCAGAACCCAGGCC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
QY 901 AAATCTCATAGATTCTCTCAGCAAGTTTTCAGAACGACGAGGACGATTTGTATGAGCAGTTCC 960
DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
QY 961 GTACCGACGACGAATTTCCCGGTCGATTTTACGCGCTTAAACCGCGGACGCGTGGATCAGG 1020
DB 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
QY 1021 GATTTGAAGACGACGATCTCAGCAAGAGCT 1050
DB 332 AspLeuLysArgProAlaGlnGlnGluAla 341

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Kozminski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 SEQUENCE FROM N.A.
 RC STRAIN-CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RC Strauberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC020041; AA020041.1; -;
 DR MGD; MGI:107438; Cab39
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39843 MW; E7FECA529D6FE811 CRC64;
 Alignment Scores:
 Pred. No.: 1.07e-112 Length: 341
 Score: 1619.50 Matches: 326
 Percent Similarity: 93.43% Conservative: 1
 Best Local Similarity: 93.14% Mismatches: 14
 Query Match: 89.82% Indels: 9
 DB: 2 Gaps: 2
 US-10-089-688-1 (1-1053) x Q8VDZ8 (1-341)
 QY 1 ATGCGCTCCGTTGGGAAGTCTCACAAATCTCCAGACAGCATTTGTGAAGATCTGAAG 60
 DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspLeuValLysAsnLeuLys 20
 QY 61 GAGAGCATGCTGTTCTGGAAGAGCAAGACATTTCTGATATAAAGCAAGAAAGCGTACA 120
 DB 21 GluSerMetAlaValLeuGluLysGlnAspLeuSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAGTTTCCAAATCTGGTCCATCAAGAAATCTGATGACCAAAATGAAAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLeuTyrGlyThrAsnGluLys 60
 QY 181 GAGCCTCAGACAGAGCATGCTCAACTTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
 QY 241 AGCACCTCGTAGCTGATTACAGCTCATTTGAGTGGGCAAGAAAGCGTGGCTCAA 300
 DB 81 GlyThrLeuValAlaAspLeuGlnLeuLeuAspPheGluGlyLysLysAspValAlaGln 100
 QY 301 ATTTTCACAAATATCTCAGACAGCAAAATTTGGTACGAGAACTCTACTGTTGATACATC 360
 DB 101 IlePheAsnAsnLeuLeuArgGlnLeuGlyThrArgThrProThrValGluTyrIle 120
 QY 361 TGACCCCAACAGAAATATTTGTTTCATGTTATGAAGGGTATGAATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
 QY 421 CTAAATTTGGGAATATTTGAAGATGTCATCAGACATGAACCATCTTGCAGAAATCAAT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
 QY 481 TTGTGGTCCGAACAGTTTATGATTTCTTCAGATATGTCGAATGTCACATTTGCATA 540

DB 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATCATTTGCCATTCACAGGATTACTTACACAGACATAAATGCTCAGTGCA 600
 DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 QY 601 GAATTTTGAACACAGCATTTATGATATGATTTTTCAGTGAATATGAGAAAGTACTTCATCA 660
 DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 QY 661 GAAAAATATGTCACAAAAGACAGTACAGTCAAGCTTCTCGTGAACCTACTACTAGATAGA 720
 DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 240
 QY 721 CACAACCTTCACAATATGACAAAATACATCAGTAAACCTCAGAACCTCAAAATTAATGATG 780
 DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 QY 781 AACCTGCTGGAGACAAAAGTCGCAACATCAGTTTGGAGCCCTTTCACGTTTTTAAGGTG 840
 DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 QY 841 TTTGTAGCCCAATCTTAACAGACGACGCCATCCTAGACATCCTCCTCAAGAACCCAGCC 900
 DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
 QY 901 AAACCTCATAGATTTCTCAGCAAGTTTTCAGAACGACGACGAGCATTTGTATGAGCAGTTCC 960
 DB 301 LysLeuLeuGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 QY 961 GTACCGACGACGAATTTCCCGGGTCGATTTACGGCTTAAACCGCGGACGCTGGGATCAGG 1020
 DB 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 QY 1021 GATTTTCAGACAGACCATCAGCTCAGCAAGAGCT 1050
 DB 332 AspLeuLysArgAlaAlaGlnGlnGluAla 341
 RESULT 3
 CB39 MOUSE STANDARD; PRT; 341 AA.
 ID_QB39_MOUSE Q06138;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Calcium binding protein 39 (Mo25 protein).
 GN Name-Cab39; Synonyms=Mo25;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93119656; PubMed=8418809;
 RA Miyamoto H., Matsushiro A., Nozaki M.;
 RT "Molecular cloning of a novel mRNA sequence expressed in cleavage
 RT stage mouse embryos";
 RL Mol. Reprod. Dev. 34:1-7(1993).
 CC -I- FUNCTION: Together with the STE20-related adaptor-alpha (STRAD
 CC alpha) pseudo kinase, forms a regulatory complex capable of
 CC stimulating the activity of Strk1.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -I- DEVELOPMENTAL STAGE: Transcribed during early mouse development.
 CC Detected at all developmental stages from the egg through the
 CC blastocyst, most abundant at the 2-cell stage.
 CC -I- SIMILARITY: Belongs to the Mo25 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; S51858; AAB24801.1; -;
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39842 MW; E7F668529D6FE811 CRC64;

Alignment Scores:

Pred. No.: 2, 52e-112 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 1 Gaps: 2

US-10-089-688-1 (1-1053) x CB39_MOUSE (1-341)

```
QY 1 ATGCGTTCCTGGTGGAGTCTCAAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGCTGTTCTGGAAGCAAGACATTTCTGATAAAGCAGAAAAGGCTACA 120
DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAAGTTTCCAAAATCTGGTTCATGAAGAAATTTCTGTATGGCACAATGAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLeuTyrglyThrAsnGluLys 60
QY 181 GAGCTCAGACAGAGCATGCTCAACTGCTCAAGACTCTATATAGTGGCTCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrglyThrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTGTGATGTTACAGCTCATTCAGTTCAGGCAAGAAAGCGTGGCTCAA 300
DB 81 GlyThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
QY 301 ATTTTCAACATATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTCTGTGATATCATC 360
DB 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrlle 120
QY 361 TGCACCCACAGATATTTGTTTCATGTTATGAAAGGCTATGAACTCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrglySerProGluIleAla 140
QY 421 CTAATTTGTGAATAATGTTAAGAGAAATGCATCAGATGAACCACTTCGAAAAATCATT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGTGGTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCAACATTTGACATA 540
DB 161 LeuTrpSerGluGlnPheTyrglyAspPheArgTyrglyValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAACTTACTTCAATCA 660
DB 201 GluPheLeuGluGlnHisTyrglyAspPheSerGluTyrglyLeuLeuLeuHisSer 220
QY 661 GAAATTTATGTCACAAAAGCAGTCACTGAGCTTCTCGTGAAGCTACTACTAGATAGA 720
DB 221 GluAsnTyrglyValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
QY 721 CACAACTTCACATATACAAAATACATCATGTAACCTGAGAACTTCAAAATTAATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrlleSerLysProGluAsnLeuLeuMetMet 260
QY 781 AACCTGCTGCAGACAAAAAGTCGCAACATCCAGTTTGAGGCTTTTACGTTTTTAAGGTG 840
```

```
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCAATCTTAACAGAGCAGCCCATCTCTAGACATCTCTCAAGAACAGGCGC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
QY 901 AAACCTCATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGAGCAGCGATTTGTATGAGCAGTTCC 960
DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
QY 961 GTACCGACGACGAATTCCTCGGTCGATTTACGCTTAAACCGCGAGCGTGGGATCAGG 1020
DB 316 AspGluGlnPheAsnAspGluLysThrTyrglyLeuValLysGln-----IleArg 331
QY 1021 GATTTGAAAGAGACAGCTCAGCAAGAGCT 1050
DB 332 AsnLeuLysArgAlaAlaGlnGlnGluAla 341

RESULT 4
Q6IP72 PRELIMINARY; PRT; 341 AA.
AC Q6IP72;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MGC78903 protein.
GN Name=MGC78903;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative."
Dev. Dyn. 225:384-391(2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S.L., Strausberg R.;
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
EMBL; BC072045; AAH72045.1; -.
DR InterPro; IPR008938; ARM.
```

```
DR InterPro: IPR004892; Mo25.
DR Pfam: PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39777 MW; 5AFEAF9325B2F16 CRC64;

Alignment Scores:
Pred. No.: 2,52e-112 Length: 341
Score: 1614.50 Matches: 323
Percent Similarity: 93.71% Conservative: 5
Best Local Similarity: 92.29% Mismatches: 13
Query Match: 89.55% Indels: 9
DB: 2 Gaps: 2

US-10-089-688-1 (1-1053) x Q6IP72 (1-341)

QY 1 ATCCGCTCCGTTGGAGTCTCAAAATCTCCAGCAGACATTTGGAAGATCTGAAG 60
DB 1 MetProPheProPheGlySerHisLeuSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTGGAAGCAGACATTTCTGATRAAAAAGCAGAAAAGGCTACA 120
DB 21 GluSerIleAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAGAAGTTTCCAAAATCTGGTTGCCATGAAGAAATTTCTGTATGCCACAAATGAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLysLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCCTCAGACAGACAGTAGTCTCAACTTGTCTCAAGAACTTATATAGTGGCTCCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGACCCCTGCTAGTGTATTTACAGCTCATTGACTTTGAGGGCAGAAAGACGCTGCTCA 300
DB 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACAATATTTCTCAGAGACAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
DB 101 IlePheAsnAsnLeuLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCAACCAACAGAAATATTTTGTTCATCTTATTAAGAGGTTATGAATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAAATTTGGTAATATTTAAGAGATGATCATGACATCAACCACTTGCAGAAATCAATT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
QY 481 TTGTGGTCGGAACAGTTTATGATTTCTCAGATATCTCGAATGTCAACATTTGCACATA 540
DB 161 LeuTrpSerGluProPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATCATTTGCCACATTCAGGATTTACTTACAAGACATAAATTTGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuSerAla 200
QY 601 GAATTTTGGACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
QY 661 GAAATTTATGTGACAAAAGACAGTCACTGAAGCTTCTCGTGAACCTACTACTAGATAGA 720
DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLysGluLeuLeuLeuAspArg 240
QY 721 CACAACCTTCAATTTATGACAAAATATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACTGCTGGCAGACAAAAGTCGCAACATCATGTTTGAGCGCTTTCACGTTTTTAAGGTG 840
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTACCCATCTTACACAGACGCGCCCATCTCTAGACATCTCTCTCAGAACCCAGGCC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProValLeuAspIleLeuLeuLysAsnGlnSer
```

Query Match: 89.55% Indels: 9 Gaps: 2
 DB: 2

US-10-089-688-1 (1-1053) x AAH72045 (1-341)

QY 1 ATGCGCTTCCGTTTGGAGTCTCAAAATCTCCAGCAGACATTGCTGAAGAACTCTGAAG 60
 DB 1 MetProPheProGlySerHisLysSerProAlaAspLeuValLysLeuLys 20
 QY 61 GAGAGCATGCTGCTTCTGGAAGAGCAGACATTCTGATAAAAAACGAGAAAGGCTACA 120
 DB 21 GluSerLeuAlaValLeuGluLysGlnAspLeuSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAAGTTCCAAAATCTGGTGGCCATGAAAGAAATCTGATGGCACAATGAAAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
 QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysThrAsnSerGlyLeuLeu 80
 QY 241 AGCACCCTGGTAGCTGATTACAGCTCAATTCAGTCAATTCAGTCAATTCAGTCAAT 300
 DB 81 GlyThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
 QY 301 ATTTTCAACATATCTCAGAGACAAATTTGGTACGAGAACTCTACTGTTGAATCATC 360
 DB 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluThrLeu 120
 QY 361 TGACCCCAACAGATATTTTCTCATGTTATTTAGAGGGTATGATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyThrGluSerProGluLeuAla 140
 QY 421 CTAATTTTGGAAATATTTAAGAGATGCAATGCAATGCAATGCAATGCAATGCAAT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
 QY 481 TTGGTTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCACATTTTGAATA 540
 DB 161 LeuTrpSerGluProPheThrAspPheArgGlyValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATGCAATTTGCCACATTTCAAGATTTACTTACAGACATAAATGCTCAGTCA 600
 DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuLeuLeuLeuLeuLeuLeu 200
 QY 601 GAATTTTGGAAACAGCATATGATGATTTTCTAGTGAATATGAGAGTTACTTACTTCA 660
 DB 201 GluPheLeuGluGlnHisThrAspArgPhePheSerGluThrGluLysLeuLeuHis 220
 QY 661 GAAATTTATGTCACAAAAGACAGTCACTGAAGCTTCTCGGTGAATCTACTACTAGATA 720
 DB 221 GluAsnThrValThrLysArgGlnSerLysLeuLeuLeuLeuLeuLeuLeuLeu 240
 QY 721 CACAACTTCACAAATATACAAAATATACATCAATGAACTGAACTCAATGAACTGAA 780
 DB 241 HisAsnPheThrIleMetThrLysThrLysSerLysProGluAsnLeuLysLeuMet 260
 QY 781 AACCTGCTCGGACAGAAAAGTCGACATCCAGTTTGGAGCTTTCAGCTTTTAAAGTG 840
 DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 QY 841 TTTGTAGCCCAATCTTAAACAGACGACCCATCTCTAGACATCTCTCTCAAGAACACG 900
 DB 281 PheValAlaAsnProAsnLysThrGlnProValLeuAspIleLeuLeuLysAsnGlnSer 300
 QY 901 AAACCTCATAGATTCCTCAGCAAGTTTCAAGACGACGAGCGGATTTGATGACGATGCC 960
 DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 QY 961 GTACCGACGACGAAATTCGCGGTTCGATTTACGGGTTAAACCGGACGCGGTGGATCAGG 1020
 DB 316 AspGluGlnPheAsnAspGluLysThrThrLeuValLysGln-----IleArg 331

QY 1021 GATTTTGAAGACAGACAGCTCAGCAAGAGCT 1050
 DB 332 AspLeuLysArgProAlaGlnGlnAspAla 341

RESULT 6
 Q724X0 PRELIMINARY; PRT; 341 AA.
 AC Q724X0;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-NAR-2004 (TRENBLrel. 26, Last annotation update)
 DE MO25-like protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhou Y., Yu L., Zhao S.Y.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF134480; AAP97257.1; --
 DR InterPro: IPR004892; Mo25.
 DR Pfam: PF03204; Mo25; 1
 DR PFam: PF03204; Mo25; 1
 SQ SEQUENCE 341 AA; 39774 MW; D3A60160E78C7A6D CRC64;

Alignment Scores:
 Pred. No.: 1.58e-110 Length: 341
 Score: 1590.50 Matches: 323
 Percent Similarity: 92.86% Conservative: 2
 Best Local Similarity: 92.29% Mismatches: 16
 Indels: 9
 Query Match: 88.21% Gaps: 2
 DB: 2

US-10-089-688-1 (1-1053) x Q724X0 (1-341)

QY 1 ATGCGCTTCCGTTTGGAGTCTCAAAATCTCCAGCAGACATTGCTGAAGAACTCTGAAG 60
 DB 1 MetProPheProGlySerHisLysSerProAlaAspLeuValLysLeuLys 20
 QY 61 GAGAGCATGCTGCTTCTGGAAGAGCAGACATTCTGATAAAAAACGAGAAAGGCTACA 120
 DB 21 GluSerMetAlaValLeuGluLysGlnAspLeuSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAAGTTTCCAAAATCTGGTGGCCATGAAAGAAATCTGATGGCACAATGAAAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
 QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysThrAsnSerGlyLeuLeu 80
 QY 241 AGCACCCTGGTAGCTGATTACAGCTCAATTCAGTCAATTCAGTCAATTCAGTCAAT 300
 DB 81 SerThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
 QY 301 ATTTTCAACATATCTCAGAGACAAATTTGGTACGAGAACTCTACTGTTGAATCATC 360
 DB 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluThrLeu 120
 QY 361 TGACCCCAACAGATATTTTCTCATGTTATTTAGAGGGTATGATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyThrGluSerProGluLeuAla 140
 QY 421 CTAATTTTGGAAATATTTAAGAGATGCAATGCAATGCAATGCAATGCAATGCAAT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
 QY 481 TTGGTTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCACATTTTGAATA 540
 DB 161 LeuTrpSerGluProPheThrAspPheArgGlyValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATGCAATTTGCCACATTTCAAGATTTACTTACAGACATAAATGCTCAGTCA 600

Db 191 AlaSerAspAlaPheAlaThrPhelysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 Qy 601 GAATTTTGGACAGCAGCAATATGATAGATATTTTCAGTGAATATGAGAGTACTTCAITCA 660
 Db 201 GlupheMetValLeuLeuSerPheArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 Qy 661 GAAATTTATGTGCAAAAAGACAGTCACTGAGCTTCTCGGTGAACACTACTACTAGATAGA 720
 Db 221 GluasentYrValThrLysArgGlnSerLeuLysLeuLeuGluGlyLeuLeuLeuAspArg 240
 Qy 721 CACAACCTTCAAAATATGACAAAATACATCATGATAAACCCTGAGAACCTCAAAATTAATGATG 780
 Db 241 HisanPheThrileMetThrLysTyrIleSerLysProGluasnLeuLysLeuMetMet 260
 Qy 781 AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTGAAGCGCTTTTCACGTTTTTAAGGTG 840
 Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 Qy 841 TTTGTAGCCATCTTACAAAGACGAGCCCATCTAGACATCTCTCAAGAACCCAGGCC 900
 Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
 Qy 901 AAATCTATAGAGTTCTCCAGCAAGTTTCAGAAACGACAGGAGCGATTGTATGAGCAGTTCC 960
 Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 Qy 961 GTACCGACGACGAATTCGCGGTGCGATTTCACGCTTAACCGCGAGCGGTGGATCAGG 1020
 Db 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 Qy 1021 GATTTGAAGACACAGCTCAGCAGAGAGCT 1050
 Db 332 AspLeuLysArgProAlaGlnGlnGluAla 341

RESULT 7
 Q61QL2 PRELIMINARY; PRT; 341 AA.

ID Q61QL2
 AC Q61QL2; 2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ZGC:86716
 GN Name=ZGC:86716;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071393; AAH71393.1; -
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39792 MW; 9CEB4A665F3C6614 CRC64;

Alignment Scores:
 Pred. No.: 1.45e-106 Length: 341
 Score: 1537.50 Matches: 306
 Percent Similarity: 91.43% Conservative: 14
 Best Local Similarity: 87.43% Mismatches: 21
 Query Match: 85.27% Indels: 9
 DB: 2 Gaps: 1

US-10-089-688-1 (1-1053) x Q61QL2 (1-341)

Qy 1 ATGCCGTTCCCGTTTGGGAAGTCTCAAAATCTCCAGACAGCATTTGTGAAGAATCTGAAG 60
 Db 1 MetProPheProPheValLysSerHisLysCysProAlaAspIleValLysAsnLeuLys 20
 Qy 61 GAGACATGCGCTGTTCTGGAAAAGACAGCATTTCTGATAAAAAGCAGAAAAGGTACAC 120
 Db 21 AspAsnMetThrileLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaSer 40
 Qy 121 GAAGAAGTTCCAAAATCTGGTTGCCATGAAGAATCTGTATGGCACAAATGAAAAA 180
 Db 41 GluGluValSerLysSerLeuLeuSerMetLysGluIleLeuTyrGlyThrAsnGluLys 60
 Qy 181 GAGCCTCAGACAGACAGTAGCTCAACTTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
 Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
 Qy 241 AGCACCTGGTAGCTGATTTCAGCTCATTCAGCTTTGAGGGCAAAAAGACGTGGCTCAA 300
 Db 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 100
 Qy 301 ATTTTCAACAATATTTCTCAGAACACAAATGGTACGAGAACTCTACTGTTGAATACATC 360
 Db 101 IlePheAsnAspIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrLeu 120
 Qy 361 TGACCCCAACAGCAATATTTGTTTCATGTTATTCAAAAGGGTATCAATCTCCAGAAAATAGCT 420
 Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProAspIleAla 140
 Qy 421 CTAATTTGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAAAATCAAT 480
 Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleThr 160
 Qy 481 TTGTGGTCCGACAGTTTTATGATTTTTCAGATATGTCGAATGTCACATTTGACATA 540
 Db 161 LeuCysSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
 Qy 541 GCTTTCAGATGTCATTTCACATTCAGAGGATTTACTTACAAGACATATAATTTGCTCAGTGA 600
 Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuSerAla 200
 Qy 601 GAATTTTGGACAGCAGATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
 Db 201 GluPheLeuGluGlnHisTyrAspPhePheSerGluTyrGluLysLeuLeuHisSer 220
 Qy 661 GAAAATATGTGCAAAAAGACAGTCACGTGAGCTTCTCGGTCAACTACTACTAGATAGA 720
 Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyLeuLeuLeuAspArg 240
 Qy 721 CACAACCTTCACAATATGACAAAATACATCAGTAAACCTCAGAACCTCAAAATTAATGATG 780
 Db 241 HisanPheThrileMetThrLysTyrIleSerLysProGluasnLeuLysLeuMetMet 260
 Qy 781 AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTGAAGCGCTTTTCACGTTTTTAAGGTG 840

```

Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTGTAGCCAACTCTAACAAGACGACGCCATCTAGACATCTCTCAAGAACAGGCCC 900
Db 281 PheValAlaAsnProAsnLysThrProProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAACATCATAGAGTTCTCCAGCAAGTTTCAGACACAGGACGCGATTGTATGACAGTTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgAlaGluAspGluGlnPheSer 320
Qy 961 GTACCGACGACGAATCCCGGGTCGATTACCGGCTTAACCGCGGACGCGTGGGATCAGG 1020
Db 321 AspGluLysThrTyr-----LeuIleLysGlnIleArg 331

Qy 1021 GATTGAAGACAGCAGTCTAGCAAGAAGCT 1050
Db 332 AspLeuLysArgProThrProGlnAspAla 341

RESULT 8
AAH71393
ID AAH71393 PRELIMINARY; PRT; 341 AA.
AC AAH71393;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC071393; AAH71393.1; -.
KW Hypothetical protein.
SQ SEQUENCE 341 AA; 39792 MW; 9CEE4A665F3C6614 CRC64;

```

Alignment Scores:

Pred. No.:	1,45e-106	Length:	341
Score:	1537.50	Matches:	306
Percent Similarity:	91.43%	Conservative:	14
Best Local Similarity:	87.43%	Mismatches:	21
Query Match:	85.27%	Indels:	9
DB:	2	Gaps:	1

```

US-10-089-688-1 (1-1053) x AAH71393 (1-341)
Qy 1 ATGCCGTTCCCGTTGGGAAGTCTCAAAATCTCCAGACAGACATTGTGAAGATCTGAAG 60
Db 1 MetProPheProPheValLysSerHisLysCysProAlaAspIleValLysAsnLeuLys 20
Qy 61 GAGACATCGCTGTTCTCGAAAAAGCAAGACATTTCTGATAAAAAACAGCAAAAGGCTACA 120
Db 21 AspAsnMetThrIleLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaSer 40
Qy 121 GAAGAAGTTTCCAAAAATCTGTTGCCATCAAGAAGAAATCTGATGGCAAAATGAAGAAA 180
Db 41 GluGluValSerLysSerLeuLeuSerMetLysGluIleLeuTyrGlyThrAsnGluLys 60
Qy 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
Qy 241 AGCACCCCTGCTAGCTGATTTACAGCTCATGCTTTGAGGGCAAAAAAGACGTGGCTCAA 300
Db 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 100
Qy 301 ATTTTCAACATATTTCTCAGACAGCAAAATGGTACGAGACTCCTACTGTTGAATACATC 360
Db 101 IlePheAsnAspIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrLeu 120
Qy 361 TGCACCCCAACAGAAATTTTGTTCATGTTATTCAAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProAspIleAla 140
Qy 421 CTAATTTGTGGAATAATGTTTAAGAGAAATGTCATCAGACATGCAACCACTTGCAGAAAATCATT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleThr 160
Qy 481 TTGTGGTCGGAACACAGTTTATGATTTCTTCAGATATGTCGAAATGCAACATTTGACATA 540
Db 161 LeuCysSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
Qy 541 GCTTCAGATCATTGTCACATTCAGAGATTACTTACAGACATCAAAATGTCAGTGCA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuSerAla 200
Qy 601 GAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
Db 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer 220
Qy 661 GAAATTTATGTCACAAAAGACAGTCACTCAAGCTTCTCGTGAACCTACTACTAGATAGA 720
Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLysAspArg 240
Qy 721 CACAACCTTCACAATTTATGACAAAATATACATCAGTAAACCTCAGAACCTCAAAATTAATGATG 780
Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
Qy 781 AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTTCAGGCTTTTTCAGTTTTCAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTTGTAGCCATCTTAACAGACGACGCCCATCTAGACATCTCTCAAGAACAGGCCC 900
Db 281 PheValAlaAsnProAsnLysThrProProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAACATCATAGAGTTCTCCAGCAAGTTTCAGAACACAGGACGCGATTGTATGAGCAGTTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgAlaGluAspGluGlnPheSer 320
Qy 961 GTACCGACGACGAATCCCGGGTCGATTACCGGCTTAACCGCGGACGCGTGGGATCAGG 1020
Db 321 AspGluLysThrTyr-----LeuIleLysGlnIleArg 331
Qy 1021 GATTGAAGACAGCAGCTCAGCAAGAAGCT 1050
Db 332 AspLeuLysArgProThrProGlnAspAla 341

```

```

RESULT 9
Q6PAB4 PRELIMINARY; PRT; 337 AA.
AC Q6PAB4;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE MGC68674 protein.
GN Name=MGC68674;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Hulyk S.W.,
RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391(2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060384; AAH60384.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 337 AA; 39183 MW; CCAB68F5D8569B57 CRC64;

Alignment Scores:
Pred. No.: 9,41e-92 Length: 337
Score: 1339.50 Matches: 268
Percent Similarity: 87.83% Conserved: 35
Best Local Similarity: 77.68% Mismatches: 29
Query Match: 74.29% Indels: 13
DB: 2 Gaps: 4

US-10-089-688-1 (1-1053) x Q6PAB4 (1-337)

QY 1 ATGCGGTTCCGGTTGGGAAGTCTCACAAATCTCCAGCAGACATTTGTGAAGATCTCAAG 60
DB 1 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysThrLeuLys 22

```

```

QY 61 GAGAGCATGGCTGTTCTGGAAAAGCAAGACACATTTCTGATATAAAAAAGCAGAAAGCGCTACA 120
DB 23 AspAsnMetalLeuLeuGluArgGln-----AspLysLysThrGluLysAlaSer 39
QY 121 GAAGAAGTTTCCAAAATATCGTTGCCATGAAGAAATTCGTATGCGCACAATGCAAAA 180
DB 40 GluGluValSerLysSerLeuGlnAlaThrLysGluIleLeuCysGlyThrGlyAspLys 59
QY 181 GAGCCTCAGACAGAAAGCAGTAGCTCAACTTCTCTCAAGAACTCTATAATAGTGGGCTCCTT 240
DB 60 GluProGlnThrGluThrValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 79
QY 241 AGCACCTGTGTAGTGTATACAGCTCATGACTTTGAGGGCAAAAAGAGCGTGCTCAA 300
DB 80 ValThrLeuIleAlaLeuLeuHisLeuIleAspPheGluGlyLysLysAspValSerGln 99
QY 301 ATTTTCAACAATATCTCAGAGACAAATGGTACGAGAACTCTCTACTGTTGAATACATC 360
DB 100 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
QY 361 TGCAACCAACAGAAATATTTGTTTCATGTTATTTGAAAGGCTATGAATCTCCAGAAATAGCT 420
DB 120 SerSerHisGlnHisIleLeuPheIleLeuLeuLysGlyTyrGluSerProGlnValAla 139
QY 421 CTAAATTTGTGAATAATGTTAAGAAATGCATCAGACATGAACCACTTGCAAAATCATT 480
DB 140 LeuHisCysGlyIleMetLeuArgGluCysValArgHisGluProLeuAlaLysValIle 159
QY 481 TTGTGGTCGACAGACAGTTTATGATTTCTTCAGATATCTCGAAATGTCACATTTTGACATA 540
DB 160 LeuTyrSerGluGlnPheGlyAspPhePheLysTyrValGluMetSerThrPheAspIle 179
QY 541 GCTTCAGATGCATTTGCCACATCAAGAGTTTACTTCAAGACATAAATTTGCTCAGTGCA 600
DB 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuMetValAla 199
QY 601 GAATTTTGAACAGCAGCATATGATAGATTTTTCAGTGAATATGAGAAAGTTACTTCAATCA 660
DB 200 GluPheLeuGluGlnAsnTyrAspArgIlePheAsnAspTyrGluLysLeuLeuHisSer 219
QY 661 GAAATTTATGTCACAAAAGACAGTCACCTCAAGCTTCTCGTGAACCTACTACTAGATAGA 720
DB 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAACCTTCAAAATATGACAAAATACATCAGTAAACCTCAGAACCTCAAAATTAATCATG 780
DB 240 HisAsnPheSerIleMetThrLysTyrIleSerLysProGluLeuLeuLysLeuMetMet 259
QY 781 AACCTGCTCGAGACAAAAGTCGCAACATCCAGTTTGAGCGCTTTTCAAGTTTAAAGGTG 840
DB 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCCCAATCTTAACAGAGCGCAGCCCATCTAGACATCTCTCTCAGAACCCAGGCC 900
DB 280 PheValAlaAsnProAsnLysThrGlnProIleValAspIleLeuLeuLysAsnGlnThr 299
QY 901 AAACTCATAGAGTTCTCTCAGCAAGTTTTCAGAACAGCAGGACGATTTGATGAGCAGTTCC 960
DB 300 LysLeuIleAspPheLeuSerSerPheGlnLysAspArgThrAsp-----Asp 315
QY 961 GTACCGACGACGAATTTCCCGGTCGATTTACGCGTTAAACCGCGAGCGCGTGATCAGG 1020
DB 316 GluGlnPheThrAspGluLysAsnTyrLeu-----IleLysGlnIleArg 330
QY 1021 GATTTCAGAGACCA 1035
DB 331 AspLeuLysLysPro 335

RESULT 10
AAH60384 PRELIMINARY; PRT; 337 AA.
ID AAH60384
AC AAH60384;
DT 02-MAR-2004 (TReMBLrel. 27, Created)

```

DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC68674 protein.
GN MGC68674.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060384; AAH60384.1; -.
SQ SEQUENCE 337 AA; 39183 MW; CCAE68F5D8569B57 CRC64;

Alignment Scores:
Pred. No.: 9.41e-92 Length: 337
Score: 1339.50 Matches: 268
Percent Similarity: 87.83% Conservative: 35
Best Local Similarity: 77.68% Mismatches: 29
Query Match: 74.29% Indels: 13
DB: 2 Gaps: 4

US-10-089-688-1 (1-1053) x AAH60384 (1-337)
QY 1 ATCGCGTTCCTGGGAGTCTCACAATCTCCAGCAGACATTTGTAAGAATCTGAAG 60
Db ||||| STANDARD; PRT; 337 AA.
4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysThrLeuLys 22
QY 61 GAGACATGGCTGTTCTGGAAAGCAAGACATTTCTGATAAAGCAGAGAAAGGCTACA 120
Db :::::|||||
23 AspAsnMetAlaLeuLeuGluArgIn-----AspLysLysThrGluLysAlaSer 39
QY 121 GAAGAAGTTCCAAAATCTGGTTCGCATGAAGAAATTTCTGTATGGCACAATAATGAAAA 180
Db |||||
40 GluGluValSerLysSerLeuGlnAlaThrLysGluIleLeuCysGlyThrGlyAspLys 59
QY 181 GAGCCTCAGACAGACAGTAGTCACTTGTCTCAAGAAGTCTATATAGTGGGCTCCTT 240

Db |||||
60 GluProGlnThrGluThrValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 79
QY |||||
241 AGCACCTGGTAGCTGATTACAGTCTCAGTCTTGGGGCAAAAACACGCTGGCTCAA 300
Db |||||
80 ValThrLeuIleAlaAsnLeuHisLeuIleAspPheGluGlyLysAspValSerGln 99
QY |||||
301 ATTTTCAACAATATTCTCAGAGACAAAATGCTAGCAGAACTCTCTACTGTTGAATACATC 360
Db |||||
100 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
QY |||||
361 TGCACCCAAACAGATATTTTGTTCATGTTTAAAGGGTATGATCTCTCCAGAAATAGCT 420
Db :::::|||||
120 SerSerHisGlnHisIleLeuPheIleLeuLeuLysGlyTyrGluSerProGlnValAla 139
QY |||||
421 CTAAATTTGGTAATATGTTAAGAGAAATCATCAGACATGAACCATCTCCAAAAATCATT 480
Db |||||
140 LeuHisCysGlyIleMetLeuArgGluCysValArgHisGluProLeuAlaLysValIle 159
QY |||||
481 TTGTGTCGGAACAGATTTTATGATTTCTTCAGATATGTCGAAATGTCACATTTGACATA 540
Db |||||
160 LeuTyrSerGluGlnPheGlyAspPhePheLysTyrValGluMetSerThrPheAspIle 179
QY |||||
541 GCTTCAGATGCTATTTCCCATTCATCAAGGATTACTTACAGACATATAAATGCTCAGTGCA 600
Db |||||
180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuMetValAla 199
QY |||||
601 GAATTTTTCGACAGCATATGATAGATATTTTTCAGTGAATATGAGAAGTACTTACTTCAATCA 660
Db |||||
200 GluPheLeuGluGlnAsnLysThrAspArgIlePheAsnAspTyrGluLysLeuHisSer 219
QY |||||
661 GAAAAATTTGTGACAAAAAGACAGTCACTCAAGCTTCTCGGTGAACACTACTACTAGATAGA 720
Db |||||
220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGlyGluLeuLeuAspArg 239
QY |||||
721 CACAACTTCAATATTATGACAAATATCATCATCAAGTAAACCTGAGAACCTCAATTAATGATG 780
Db |||||
240 HisAsnPheSerIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY |||||
781 AACCTGTCGGACACAAAAGTCGCAACATCCAGTTTGGAGCCCTTTCACGTTTTTAAGGTG 840
Db |||||
260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY |||||
841 TTTGTAGCCAAATCTTAACAGACGCGCCATCTCTAGACATCTCTCTCAAGAACCGGCC 900
Db |||||
280 PheValAlaAsnProAsnLysThrGlnProIleValAspIleLeuLeuLysAsnGlnThr 299
QY |||||
901 AAATCTATAGATTTCTCAGCAAGTTTTCAGAACGAGGACGGATGTTATGAGCAGTTCC 960
Db |||||
300 LysLeuIleAspPheLeuSerSerPheGlnLysAspArgThrAsp-----Asp 315
QY |||||
961 GTACCGACGACGAATTCCTCGGTCGATTTACGGCTTAAACCGCGGACGCGTGGGATCAGG 1020
Db |||||
316 GluGlnPheThrAspGluLysAsnTyrLeu-----IleLysGlnIleArg 330
QY |||||
1021 GATTGAAGACACCA 1035
Db |||||
331 AspLeuLysLysPro 335

RESULT 11
C39L MOUSE
ID C39L MOUSE STANDARD; PRT; 337 AA.
AC Q9DBI6; Q8B052; Q91WB8; Q91YL0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39-like (Mo25-like protein).
GN Name=Cab39l;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

Db 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
Qy 661 GAAATTTATGTGACAAAGACAGCTCACTAGCTTCTCGTGAACCTACTACTAGTAGA 720
Db 220 GluAsnTyrValThrLysA-gGlnSerLeuLysLeuGluGlnLeuLeuAspArg 239
Qy 721 CACAACCTTCAATATTATGACAAATATACATCATGTAATACCTGAGAACCTCAAAATTAATGATG 780
Db 240 HisasnPhethrIleMetThrLysTyrIleSerLysProGluasnLeuLysLeuMetMet 259
Qy 781 AACCTGTCGAGACAAAGTCGCAATCATCAGTTTCAGCGCTTTCAGTTTAAAGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
Qy 841 TTTTACGCAATCTTACACAGACCGACCCATCTGACATCTCTCTCAAGACACCGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluLeuLeuLysLysAsnGlnPro 299
Qy 901 AAACCTATAGATTCCTCAGCAAGTTTCAGAACACAGGACGAGTTCATGAGCAGTTC 960
Db 300 LysLeuLeuGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
Qy 961 GTACCGACGAGAAATCCCGGTTCGATTGACGCTTAAACCGGACGCGTGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuLeuLysGlnIleArg 330
Qy 1021 GATTTGAGACACGAGCT 1038
Db 331 AspLeuLysLysAlaAla 336
RESULT 12
C39L HUMAN STANDARD; PRT; 337 AA.
ID Q9H9S4; Q9G7L1; Q96FG1; Q9BZ33;
AC 16-OCT-2001 (Rel. 40, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39-like (Mo25-like protein) (Antigen MLAA-34).
GN Name=CAB39L;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen G., Zhang W.G., Fu J., Cao X.M., Chen Y.X., Zhao W.H., Zhao A.Z.,
RA Han Y.H., Li F.Y., Liu X.P., Yao L.B.;
RA "Serological identification of immunogenic antigens in acute monocytic
RA leukemia".
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
SEQUENCE FROM N.A.
RN PubMed=15057823; DOI=10.1038/nature02379;
RA Dunham A., Matthews L.H., Burton D., Ashurst J.L., Howe K.L.,
RA Ashcroft K.J., Beare D.M., Burton J.C., Hunt S.E.,
RA Griffiths-Jones S., Jones M.C., Keenan S.J., Oliver K., Scott C.E.,
RA Alnsough R., Almeida J.P., Ambrose K.D., Andrews D.T.,
RA Ashwell R.I.S., Babbage A.K., Baggeley C.L., Bailey J., Bannerjee R.,
RA Barlow K.F., Bates K., Beasley H., Bird C.P., Bray-Allen S.,
RA Brown A.J., Brown J.Y., Burrill W., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M.E., Clark S.Y., Clarke G., Cline C.W.,
RA Clegg S.C., Cobley V., Collins J.E., Corby N., Coville G.J.,
RA Deloukas P., Dhami P., Dunham I., Dunn M., Earthrowl M.E.,
RA Ellington A.G., Faulkner L., Frankish A.G., Frankland J., French L.,
RA Garner P., Garnett J., Gilbert J.G.R., Gilson C.J., Ghori J.,
RA Grafham D.V., Gribble S.M., Griffiths C., Hall R.E., Hammond S.,
RA Harley J.L., Hart E.A., Heath P.D., Howden P.J., Huckle E.J.,
RA Hunt P.J., Hunt A.R., Johnson C., Johnson D., Kay M., Kimberley A.M.,
RA King A., Laird G.K., Langford C.J., Lawlor S., Leongamornlert D.A.,
RA Lloyd D.M., Lloyd C., Loveland J.E., Lovell J., Martin S.,
RA Mashreghi-Mohammadi M., McLaren S.J., McMurray A., Milne S.,
RA Moore M.F., Nickerson T., Palmer S.A., Pearce A.V., Peck A.I.,
RA Pelan S., Phillimore B., Porter K.M., Rice C.M., Searle S.,
RA Selva H.K., Showmore R., Skuce C.D., Smith M., Steward C.A.,
RA Sycamore N., Tester J., Thomas D.W., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallis J.M., West A.P., Whitehead S.L., Willey D.L.,
RA Wilming L., Wray P.W., Wright M.W., Young L., Coulson A., Durbin R.,
RA Hubbard T., Sulston J.E., Beck S., Bentley D.R., Rogers J., Ross M.T.;
RA "The DNA sequence and analysis of human chromosome 13";
RL Nature 428:522-528(2004).
[4]
SEQUENCE OF 7-337 FROM N.A.
RN PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki I., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotochi T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigetaki K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RA "Complete sequencing and characterization of 21,243 full-length human
RA cDNAs";
RL Nat. Genet. 36:40-45(2004).
CC -[1] SIMILARITY: Belongs to the Mo25 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch).

Db	200	AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer	219
Qy	661	GAATAATTATGTGACAAAAAGACAGCTACTGAAGCTTCTCGGTGAACCTACTACTAGATAGA	720
Db	220	GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg	239
Qy	721	CACAACCTTCACAATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG	780
Db	240	HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet	259
Qy	781	AACCTGCTGGAGACAAAAGTCGCAACATCCAGTTTGAGGCGCTTCACGTTTTTAAGGTG	840
Db	260	AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal	279
Qy	841	TTTGTGACCAATCCTAACAGACGACGCCCATCTCTAGACATCCTCCCTCAAGAACCCAGGCC	900
Db	280	PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro	299
Qy	901	AAACTCATAGAGTTCCCTCAGCAAGTTTCAGAACACACAGCGAGTTGTATGACGAGTTCC	960
Db	300	LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp	314
Qy	961	GTACCGACGACGAATCCCGGGTCGATTTACGCGTTAAACCGCGACGCGTGGGATCAGG	1020
Db	315	-----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg	330
Qy	1021	GATTTGAAGACACCGCT	1038
Db	331	AspLeuLysLysThrAla	336

RESULT 13		
AAQ93064	PRELIMINARY;	PRT; 337 AA.
ID	AAQ93064	
AC	AAQ93064;	
DT	31-MAR-2004 (TrEMBLrel. 27, Created)	
DT	31-MAR-2004 (TrEMBLrel. 27, Last sequence update)	
DT	31-MAR-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Antigen MAA-34.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Chen G., Zhang W.G., Fu J., Cao X.M., Chen Y.X., Zhao W.H., Zhao A.Z.,	
RA	Han Y.H., Li F.Y., Liu X.P., Yao L.B.;	
RT	"Serological identification of immunogenic antigens in acute monocytic	
RT	leukemia.";	
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AY288977; AAQ93064.1; -.	
SQ	SEQUENCE 337 AA; 39029 MW; B8FIAB5C7905EC93 CRC64;	
Alignment Scores:		
Pred. No.:	7.31e-88	337
Score:	1287.50	261
Percent Similarity:	84.97%	33
Best Local Similarity:	75.43%	39
Query Match:	71.41%	13
DB:	2	3
US-10-089-688-1 (1-1053) x AAQ93064 (1-337)		
Qy	1 ATCCGPTCCGGTTTGGGAGTCTCAAAATCTCCAGCAGACATTTGGAAGAACTCTGAAG	60
Db	4 McpProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys	22
Qy	61 GAGAGAGTGCTGTCTTGAAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAAGCTACA	120
Db	23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer	39
Qy	121 GAGAGAGTTCCAAAATCTGGTTGCCATGAAGAGAAATCTTGTAATGGCAAAATGAAAAA	180
Db	40 GluGluValSerLysSerLysLeuGlnAlaMetLysGluIleLeuLysGlyThrAsnAspLys	59

```
QY 181 GAGCTCAGACAGAGCAGTAGCTCAACTGCTCAAGAACTCTATATAGTGGCTCCTT 240
D 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuTyrSerSerGlyLeuLeu 79
QY 241 AGCACTCGTGTAGCTGATTACAGCTCACTGCTTTCAGGCGCAAAAGAGCTGGCTCAA 300
D 80 ValThrLeuLeuAlaAspLeuGlnLeuLeuAlaAspPheGluGlyLysLysAspValThrGln 99
QY 301 ATTTTCAACATATCTCAGAGAGCAAAATTTGGTACGAGAACTCTCTACTGTTGATACATC 360
D 100 IlePheAsnAsnIleLeuValArgGlnIleGlyThrArgCysProThrAlaGluTyrIle 119
QY 361 TGCACCCCAACAGATATTTTGTTCATGTTTAAAGCGGTATGAATCTCCAGAAATAGCT 420
D 120 SerSerHisProHisIleLeuValMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
QY 421 CTTAATTTGTGAATAATGTTAAAGAGATGTCATCAGACATGAACCACTTGCAGAAATCATT 480
D 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
QY 481 TTGTGCTCGGACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCACATTTGACATA 540
D 160 LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
QY 541 GCTTCAGATCATTTGCCACATTCACAGATTTACTTCAAGACATCAATGCTCAGTGCA 600
D 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY 601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATGAGAATGACTTACTTCATCA 660
D 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
QY 661 GAAATTTATGACAAAGAGACAGTCACTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
D 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAACTTCACAATATGACAAAATACATCATGATAAATCTGAGAACCTCAATTAATCATG 780
D 240 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACCTGCTGCGAGACAAAGTCGCAACATCAGTTTGGAGCTTTTCACGTTTTTAAGGTG 840
D 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCAATCTTAACAGACGCGCCATCTAGACATCTCTCTCAAGAACCCAGGCC 900
D 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
QY 901 AAATCTAGAGTTCTTCAGCAAGTTTTCAGAACGACGACGAGCGGATTGATGAGCAGTCC 960
D 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY 961 GTACCGACGCAATTTCCCGGCTCGATTTCACGCGTTAAACGCGACGCGTGGATCAGG 1020
D 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
QY 1021 GATTGAAGACGACGCT 1038
D 331 AspLeuLysLysAlaAla 336
RESULT 14
Q803V8 PRELIMINARY; PRT; 343 AA.
AC Q803V8;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Zgc:55451.
GN Name=zgc:55451;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
```

```
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC04172; AAH4172.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 343 AA; 39820 MW; F10450DA0446268A CRC64;

Alignment Scores:
Pred. No.: 7,97e-88 Length: 343
Score: 1287.00 Matches: 253
Percent Similarity: 83.33% Conservative: 42
Best Local Similarity: 71.47% Mismatches: 41
Query Match: 71.38% Indels: 18
DB: 2 Gaps: 3

US-10-089-688-1 (1-1053) x Q803V8 (1-343)
QY 1 ATGCCGTTCCGTTGGGAGAGTCTCACAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
D 1 MetProPheProPheGluLysSerGlnLysSerProAlaGluIleValLysSerLeuLys 20
QY 61 GAGAGCATGCTGTTCTGGAAAAGCAGACATTTCTCAT--AAAAAGCAGAAAAGGCT 117
D 21 GluAsnValAlaTyrLeuGluLysLeuGluSerSerGluSerLysLysCysGluLysVal 40
QY 118 ACAGAAGAAGTTTCCAAAAATCTGTTGCCATGAAGAAATTTCTGTATGGCACAATCAA 177
D 41 AlaGluGluValSerLysAsnLeuSerSerLeuLysGluValLeuCysGlyThrGlyAsp 60
QY 178 AAAGAGCTCAGACAGACAGAGTAGCTCAACTTCTCAAGAACTCTATATAGTGGGCTC 237
D 61 LysGluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnThrAsnLeu 80
QY 238 CTTAGCACCTGGTAGCTGATTACAGTCACTTTCAGTGGGCAAAAAGAGCTGGCT 297
D 81 LeuLeuSerLeuIleAlaAsnLeuGlnArgIleAspPheGluGlyLysLysAspValVal 100
QY 298 CAAATTTTCAACAATATTCTCAGAAAGCAAAATTTGGTACGAGAACTCTCTACTGTTGAATAC 357
D 101 HisLeuPheSerAsnIleValArgArgGlnIleGlyAlaArgThrProThrValGluTyr 120
QY 358 ATCTGACCCCAACAGATATTTTCTTATGTTTGAAGGTTATGATCTCCAGAAATA 417
D 358 ATCTGACCCCAACAGATATTTTCTTATGTTTGAAGGTTATGATCTCCAGAAATA 417
```

```

Db 121 ileSerSerHisSerGlnIleLeuPheMetLeuLeuLysGlyTyrGluThrSerGluVal 140
Qy 418 GCTCTAAATTTGGAATAATTTAAGAGATGCAATGACATGAAACCACTTGCAGAAATC 477
Db 141 AlaLeuAsnGlyMetMetLeuArgGluCysLeuArgHisAspProLeuAlaArgile 160
Qy 478 ATTTTGTGGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAGCAATTTGAC 537
Db 161 ValLeuPheSerGluAspPheTyrCysPhePheArgTyrValGluMetSerThrPheAsp 180
Qy 538 ATAGCTTCAGATGATTTGCCACATTTCAAGATTTTACTTCAAGACATATAATTTCTCAGT 597
Db 181 ileAlaSerAspAlaPheAlaSerPheLysAspLeuLeuThrArgHisLysIleMetCys 200
Qy 598 GCAGAAATTTTGGAAACAGCATTTATGATATTTTTCAGTGAATATGAGAAATTTTCAAT 657
Db 201 AlaAspPheLeuGluThrAsnTyrAspArgValPheThrGluTyrGluLysLeuLeuHis 220
Qy 658 TCAGAAATTTATGACAAAGACAGTCATCTGAGCTTCTCGTGAACTACTACTACTAGAT 717
Db 221 SerGluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAsp 240
Qy 718 AGACAACTTCACAAATTTATGACAAATATACATCATGATAAATCTGAGAACCTCAAAATTAATG 777
Db 241 ArgHisAsnPheThrValAlaThrLysTyrIleSerArgAlaGluAsnLeuLysLeuMet 260
Qy 778 ATGAACCTGCTGCAGACAAAGTCGCAACATTCAGATTTTGGAGCTTTCAGTTTAAAG 837
Db 261 MetAsnMetLeuArgAspAsnSerArgAsnIleGlnPheGluAlaPheHisValPheLys 280
Qy 838 GTGTTTGTAGCAATCTTCAACAGCAGCCATCTAGACATCTCTCAGCACTCTCAGAACAG 897
Db 281 ValPheValAlaAsnProAsnLysThrGlnProValLeuAspIleLeuLeuLysAsnGln 300
Qy 945 GCCAAATCTATAGATTTCTTCAGCAAGTTTTCAGAACGACGACGAGT----- 945
Db 301 SerLysLeuValGluPheLeuSerHisPheGlnThrAspArgSerGluAspGluGlnPhe 320
Qy 946 TGTATGAGCATTCCTGACGACGACGAAATTCCTCGGTTCGATTTACCGTTAAACCCCGG 1005
Db 321 CysAspGluLysAsnTyr-----Leu 327

Qy 1006 ACGGTGGGATCAGGATTTTGAAGACGACCATCTGACGACGAA 1047
Db 328 IleLysGlnIleArgAspLeuLysArgProAlaProAlaGlu 341

RESULT 15
Q7PF07 PRELIMINARY; PRT; 332 AA.
AC Q7PF07;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000023148 (Fragment).
GN Name=ENSANGG0000021597;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01008804; EAA45510.1; -.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
FT NON TER 332 332
SQ SEQUENCE 332 AA; 38665 MW; E315A6819F3BDA67 CRC64;

```

```

Alignment Scores: 1.38e-78 Length: 332
Pred. No.: 1163.50 Matches: 226
Score: 86.08% Conservative: 46
Percent Similarity: 71.52% Mismatches: 39
Best Local Similarity: 64.53% Indels: 5
Query Match: 2 Gaps: 3
DB: 1-1053 x Q7PF07 (1-332)

Qy 1 ATCCGCTTCCCGTTGGGAAAGTCTCACAAATCTCCAGCAGACATTTGTGAAGAAATCTGAAG 60
Db 1 MetProLeu---PheGlyLysSerGlnLysSerProGlnGluLeuValLysSerLeuLys 19
Qy 61 GAGACATGCTGCTGTTCTGAAAAGCAAGACATTTCTCATATAAAAGCAGAAAAGCTACA 120
Db 20 GluAlaValAlaAsnAlaLeuGluArg-----GlyAspLysLysAlaGluLysAlaGln 36
Qy 121 GAAGAAATTTCCAAAATCTGTTGCCATGAAGAAATTTCTGTATGCCACAAATGAAGAAA 180
Db 37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuCysGlyThrAlaAspAla 56
Qy 181 GAGCCTCAGACAGAA---GCAGTAGCTCAACTTCTCAAGAACTCTTAATAGTGGGCTC 237
Db 57 GluProGlnThrGluLysValIleSerGlnLeuAlaHisGluLeuTyrSerThrAspLeu 76
Qy 238 CTTAGCACCTGCTGATGATTTACAGCTCATTCAGCTTTTTCAGGCAAAAGACGCTGCT 297
Db 77 LeuLeuLeuLeuIleGlnAsnLeuAsnArgIleAspPheGluGlyLysLysAspValAla 96
Qy 298 CAAATTTTCAACATATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTACTGTTGAATAC 357
Db 97 HisIlePheAsnAsnValIleArgArgGlnIleGlyThrArgLeuProThrValGluTyr 116
Qy 358 ATCTGCACCCCAACAGAAATTTTGTTCATGCTTATTGAAAGGCTATGAATCTCCAGAAATA 417
Db 117 IleCysThrLysProGluLysLeuPheThrLeuMetAlaGlyTyrGluHisGlnGluLys 136
Qy 418 GCTCTAAATTTGGAATAATTTAAGAGAAATGCAATGCAATGCAATGCAATGCAATGCAAT 477
Db 137 AlaLeuAsnGlyIleMetLeuArgGluCysAlaArgHisGluAlaLeuAlaLysIle 156
Qy 478 ATTTTGTGGTCGGAACAGTTTATGATTTTCTCAGATATGTCGAAATGTCAGAAATTTGAC 537
Db 157 MetLeuHisSerGluGluPhePheAsnPhePheArgTyrValGluValSerThrPheAsp 176
Qy 538 ATAGCTTCAGATGCTTTCGACATTTCAAGGATTTTCTTCAAGACATATAATTTCTCAGT 597
Db 177 IleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeuLeuSer 196
Qy 598 GCAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAATTTTCTCAT 657
Db 197 AlaGluPheLeuGluGlnAsnTyrAspLysValPheSerArgTyrGluAlaLeuLeuAsn 216
Qy 658 TCAGAAATTTATGTCAGAAAGCAAGTCTCACTGAGCTTCTCGTGAACCTACTACTAGAT 717
Db 217 SerGluAsnTyrValThrArgArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAsp 236
Qy 718 AGACAACTTCACAAATTTATGACAAATATACATCAGTAAACCTCAGAACCTCAAAATTAATG 777
Db 237 ArgHisAsnPheThrValMetThrLysTyrIleSerAsnProAspAsnLeuLysLeuMet 256
Qy 778 ATGAACCTGCTGCAGACAAAGTTCGCAACATCCAGTTTTCAGGCTTTTTCAGTTTAAAG 837
Db 257 MetAsnMetLeuLysGluThrSerArgAsnIleGlnPheGluAlaPheHisValPheLys 276
Qy 838 GTGTTTGTAGCCAAATCTTAAACAGACGACCCCATCTAGACATCTCTCCTCAGAACAG 897
Db 277 ValPheValAlaAsnProAsnLysProLysProIleMetAspIleLeuLeuArgAsnGln 296
Qy 898 GCCAAATCTATAGATTTCTTCAGCAAGTTTTCAGAACGACGACGAGT 945
Db 297 GluLysLeuValAspPheLeuThrArgPheHisThrAspArgSerGlu 312

```

Search completed: November 10, 2004, 20:51:01
Job time : 247.5 secs

Mon Nov 15 15:11:57 2004

Db	1453	AspGlnAlaLysGluIleSerTyrLeuLysAspGluIleLysLysIleSerMetLeuTyr	1472
Qy	166	GGC-----ACAATGAAAGAGCCTCAGACAGACGAGTACTCAACTTGTCTCAAGAA	219
Db	1473	GlyGluLeuLeuAsnArgLysAsnSerTyrAspGluLysValLysAsnLeuThrAsnGlu	1492
Qy	220	CTCTATAATAGTGGCTCTTAGCACCCCTGGTACTGATTACAGCTCATTGACTTTGAG	279
Db	1493	Leu-----LysGluLeuLysIleArgAsnLysLys	1502
Qy	280	GGCAAAAAGACGTGGCTCAATTT-----	303
Db	1503	GlyGluGluAlaIleAlaGluLeuAsnLysLeuLysAsnIleLysGluLysAsnLysSer	1522
Qy	304	-----TTCAACAATATCTCAGAACACAAATTTGTACGAGA	339
Db	1523	VallysGlnAsnAspGluSerSerAsnAsnIleIleThrLysAsp--GlyAspLys	1541
Qy	340	ACTCCTACTGTTGAATACATCTGCACCCACACAGATATT-----TTGTTC	384
Db	1542	ThrPro-----GluTyrValSerAsnAspLysIleGlnLysAspTrpLysAlaAsn	1559
Qy	385	ATGTATTGAAGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGGAATAATGTTAAGA	444
Db	1560	LeuValLysLeuLysGluLysGluLysProAspTrpAspIleAsnSerLeuGluLys	1579
Qy	445	GAATGCATCAGACATGAACCACTTGCAAAA-----ATCATTGTTGGTGGACACATTT	498
Db	1580	GluAsnPheArgValMetSerIleValLysGluAsnLysAsnValGlnAsnAspLysIle	1599
Qy	499	TATGATTCTTCAGATATGTCGAATGTCACATTTTGACATAGTTGCAGATGCTATTGGCC	558
Db	1600	ValGlyIleTyrSerTyrPheLysLysCysGluLysGluLysAsnAspMetLeuVal	1619
Qy	559	ACATTCAAGATTACTTACAGACATATAATTTGCTCAGTCAGATTTTTCGAAACAGCAT	618
Db	1620	Ile-----CysLeuValLysLysAspIleLeuSerIleLeuPheLeuAsnAspAsn	1636
Qy	619	TATGATAGATTTTTCAGTGAATATCAGAACTTACTTATTCAGAAAAATATGACAAA	678
Db	1637	PheValAsnLeuPheGluLysIleAspLysIleLeuTrpLysGlnMetTyrIleProThr	1656
Qy	679	-----AGACAGTCACTG	690
Db	1657	GluIleArgIleLeuPheLeuArgTyrPheSerPheLeuAspLysLeuArgAsnTyrVal	1676
Qy	691	AAGCTTCTCGTGAACCTACTACTAGAT-----AGACAACTTCCAAATTAATG	738
Db	1677	LysCysValAsnGluGluTyrValAsnAsnGluArgTyrGluTyrSerTrpAlaLeuPhe	1696
Qy	739	ACAAATATCATCGTAAACCTCAGACCTCAATTAATGATG	780
Db	1697	GlnThrTyrLeuGluThrAlaSerAsnLeuLysLysGluMet	1710
RESULT 2			
US-10-732-923-14011			
; GENERAL INFORMATION:			
; APPLICANT: Edgerton, Michael D			
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES			
; FILE REFERENCE: 38-15(52796)C			
; CURRENT APPLICATION NUMBER: US/10/732,923			
; CURRENT FILING DATE: 2003-12-10			
; PRIOR APPLICATION NUMBER: 10/310,154			
; PRIOR FILING DATE: 2002-12-04			
; NUMBER OF SEQ ID NOS: 24149			
; SEQ ID NO 14011			
; LENGTH: 4106			
; TYPE: PRT			
; ORGANISM: Equus caballus			
US-10-732-923-14011			
Alignment Scores:			
Pred. No.:	0.026	Length:	4106
Score:	106.00	Matches:	92
Percent Similarity:	35.16%	Conservative:	81
Best Local Similarity:	18.70%	Mismatches:	141
Query Match:	5.88%	Indels:	178
DB:	6	Gaps:	25
US-10-089-688-1 (1-1053) x US-10-732-923-14011 (1-4106)			
Qy	13	TTTGGGAAGTCTCAAAATCT---CCACGACATGTTGTGAAGAATCTG-	57
Db	2679	PheGlyLysLysArgLeuLeuProGlyAspGluValAspAsnLysAlaLysGlyThr	2698
Qy	58	-----AAGGAGAGATCGCTGTTCTGGAAAAGCAAGACATTTCTGTATAAAAA-	105
Db	2699	AspAsnArgAlaGluIleLeuArgLeuArgArgPheLeuLysAspArgGluLysLeu	2718
Qy	106	-----CCAGAAAAGGCTACAGAGAAGATTTCCTCAAAAT	138
Db	2719	SerLeuIleTyrAlaArgLysGlyValAlaGluGlnLysArgGluLysGluLysSer	2738
Qy	139	CTGGTTGCCATCAAGAA-ATTCGTATGGCACAATGCAAAAGAG---	183
Db	2739	GluLeuLysMetLysHisAspAlaGlnValIleLeuTyrArgSerTyrArgGlnGlyAsp	2758
Qy	184	-----CCTCAGACAGACAGCAGTCTCAA	207
Db	2759	LeuProAspIleGlnIleLysTyrSerSerLeuIleThrProLeuGlnAlaValAlaGln	2778
Qy	208	-----CTTGCTCAAGACTCTATAATAGTGGGCTCTTAGCACCCTGGTAGCT	255
Db	2779	ArgAspProIleIleAlaLysGlnLysPheGlySer---LeuPheSerGlyIleIleLys	2797
Qy	256	GATTTCACAGCTCATTCACCTTCAGGGGCAAAAAGAC-----GTGGCTCAA	300
Db	2798	GluMetAspLysTyrLysThrMetSerGluLysAsnAsnIleThrGlnLysLeuLeuGln	2817
Qy	301	ATTTTCAACAATATCTCAGACAGCAAAATGGTACGAGAACTCCTACTGTT-----	351
Db	2818	AspPheAsnAsnPheLeuAsnThrValSerPhePheProProPheIleSerCysIle	2837
Qy	352	---GAATACATCTGCCCAACAGATATTTTTCATGTTATTGAAGGGTATGATCT	408
Db	2838	GlnGluLysSerCysGlnHisAlaAspLeuLeuSerLeu-----AspPro	2852
Qy	409	CCAGAAATAGCTCTAAATTTCT-----GGAATA---ATGTTA	441
Db	2853	AlaSerValSerAlaSerCysLeuAlaSerLeuGlnGlnProValGlyValArgLeuLeu	2872
Qy	442	AGAGATGCATCAGACAT-----GAACCACTTCGCAAAATCATTTTGTGTCG	489
Db	2873	GluGluAlaLeuLeuHisLeuLeuProGluGluProProAlaLysArgValArgGlyArg	2892
Qy	490	GAACAGTTTAT---GATTTCTTCAGATATGCGAAATGTCA-----	528
Db	2893	ProCysLeuTyrProAspPheValArgTipMetGluLeuAlaLysLeuTyrArgSerIle	2912
Qy	529	---ACATTGACATAGCT-----	543
Db	2913	GlyGluTyrAspIleLeuArgGlyIlePheAsnSerGluLeuGlyThrLysGlnValThr	2932
Qy	544	-----TCAGATGCATTTGCCACATTCOAAG	567
Db	2933	GlnAsnAlaLeuLeuAlaGluAlaArgAsnAspTyrSerGluAlaValLysGlnTyrAsn	2952
Qy	568	GATTTCATTACAAGACATAAATTCCTCAGTCGACAGATTTTGGACAGCATTTATGATAGA	627
Db	2953	GluAlaLeuAsnLysGlnAspTyrValAspGlyGluProMetGluAlaGluLysAspPhe	2972
Qy	628	TTT-----TTCACTGAATATGAAAGTTA-----	651
Db	2973	TrpGluLeuAlaSerLeuAspCysTyrAsnGlnLeuAlaGluTrpLysSerLeuAlaTyr	2992

```
QY 651 ----- 651
Db 2993 CysSerThrValSerValAspSerAlaAsnProProAspLeuAenLysMetTrpAsnGlu 3012
QY 652 ---CTTCATTAGAAATATGTGACAAAA-----AGACAGTCACCTCAAGCTTCTCGGT 702
Db 3013 ProPheTyrGlnGluThrTyrLeuProTyrMetIleArgSerLysLeuLysLeuLeu 3032
QY 703 GAATCTACTAGTAGACACAACTTCACAAATATGACAAATATACATCAGTAACCTGAG 762
Db 3033 GlnGlyGluGlyAspGlnSerLeuLeuThrPheIleAspGluAlaValSerLys---Glu 3051
QY 763 AACCTCAAAATTAAGTGAACCTG-----CTG 789
Db 3052 LeuGlnLysValLeuValGluLeuHieTyrSerGlnGluLeuSerLeuLeuTyrIleLeu 3071
QY 790 CGACACAAAATGTCACACATCCAGTTTCAGGCTTTTCACGTTTTTAAGGTGTGTGTAGCC 849
Db 3072 GlnAspAspValAspArgAlaLysTyrTyrIleGluAenCysIleArgIlePheMet--- 3090
QY 850 AATCCTAACAGAGCGACGCCATCTAGACATCTCTCAAGAACCCAG----- 897
Db 3091 -----GlnSerTyrSerSerIleAspValLeuLeuGluArgSerArgLeuThrLys 3107
QY 898 -----GCCAACTCATAGATTCTCTCAGCAAGTTT 927
Db 3108 LeuGlnSerLeuGlnAlaLeuIleGluIleGlnGluPheIleSerPheIleArgLysGln 3127
QY 928 CAGAACGACGAGCGGATTGTATGACGAGTTCCTGACCGAGC----- 969
Db 3128 GlyAsn-----LeuSerSerGlnIleProLeuLysArgLeuLeuLysThr 3142
QY 970 ---ACGAATCCCGGTGCGATTACGGCTTAAACCG 1002
Db 3143 TriThrAsnArgTyrProAspAlaLysMetAspPro 3154
```

RESULT 3

```
US-10-732-923-3091
; Sequence 3091, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3091
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-3091
```

```
Alignment Scores:
Pred. No.: 0.0194 Length: 793
Score: 105.50 Matches: 75
Percent Similarity: 37.15% Conservative: 45
Best Local Similarity: 23.22% Mismatches: 132
Query Match: 5.85% Indels: 71
DB: 6 Gaps: 15
```

US-10-089-688-1 (1-1053) x US-10-732-923-3091 (1-793)

```
QY 172 AATGAAAAGAGCTCAGACAGACAGTAGCTCAACTTGTCTCAAGACTCTATATAGT 231
Db 139 AspGlnProSerProSerSerHisArgValSerAspPheSerAspGlu---TyrSerPro 157
QY 232 GGGCTCTCTAGCACCCCTGTAGCTGATTACAGCTCATCTACTTGTAGGGCAAAAAGAC 291
Db 158 SerLysValArgIleLeuValAlaAsp-----AspAspAsp 169
QY 292 GTG---GCTCAAAATTTCAACAATATTTCTCAGAAGA-----CAAATT 330
```

```
Db 170 ValGlnArgLysIleLeuAenAsnLeuLeuLysPheHisTyrAsnValThrLeuVal 189
QY 331 GGTACGAGAACTCTACTGTTGTAATACATCTGCACCCCAACAGAAATATTTGTTTCATGTTA 390
Db 190 ProAenGlyGluIleAlaTrpGluTyrIleAsnLysGlyGlnGlnLysTyrAspLeuVal 209
QY 391 TTGAAAGGGTATGATCTCCAGAAATAGCTCTAAATTTGTGGAATAATGTTAAGAGAATGC 450
Db 210 LeuThrAspValMetMetProHisIleThr-----GlyPheAspLeuLeuGlnArg 226
QY 451 ATCAGACATCAACACCTTCGCAAAA-----ATCATTTTGTGGTCGGAACAGTTTAT 501
Db 227 IleAsnAspHisProValHisArgHisIleProValIleLeuMetSerGlyThrAlaVal 246
QY 502 GATTTCCTCAGATATGTCGAAATGTCACATATTTGACATAGCTTCAGATGCTATTTGCCACA 561
Db 247 Asp---TyrLysTyrAlaAsn---AspThrIleLysIleGlyGlyGlnAspPheLeuThr 264
QY 562 TTCAAGGATTTACTTACACACATATAATCTCAGTCAGCCAGAAATTTTGGACACAGCATAT 621
Db 265 -----LysProIleAlaLysGluLeuLeuLysLysLysIle 276
QY 622 GAT-----AGATTTTTCAGTGAATATGAACTTACTT 654
Db 277 AspThrValLeuGlnSerIleTrpGlnArgLysGluGluGluTyrLysAlaThrLeu 296
QY 655 CATTCAGAAAATATGTCACAAAAGACAGCTCACTGAAGCTTCTCGGTGAACCTACTACTACTA 714
Db 297 AlaGlnGlu-----ArgGluLysGlyAsnLysLeuAlaLysGluMetGluLeu 312
QY 715 GATAGACACAACTTCACAAATATGACAAAATATGACAAATATGAACTGAGAACCTC----- 768
Db 313 LysGluHisGluIleGluGluLeuThrLysLysValSerLysMetSerSerIleSerLys 332
QY 769 AATTAATGATGAACCTGCTCGCAGACAAAAGTCGCAACATCCAG----- 813
Db 333 GluAlaMetGluSerProLeuValSerValThrArgAsnIleGluGluLeuLeuLysGln 352
QY 814 -----TTTGAGGCTTTTCACGTTTTTAAGGTGTTTGTAGCAACTCTTAAC 858
Db 353 SerSerTrpSerHisTyrGluSer-----GluIleLysGlu 364
QY 859 AAGACGACGCCCATCTCTAGACATCTCTCCTCAAGAACCCAGCCCAAACTCATAGAGTTCTCTC 918
Db 365 LysLeuSerSerIleLeuLysGluLeuGlySerSerAsnIleTyrArgProSerPheGlu 384
QY 919 AGCAAGTTTCAGAACGACGACGAGCGGATTGTATGACGAGTTCGGTA----- 963
Db 385 LysLeuIleLysAsnAspSerValAspProValThrLysSerPheLeuValSerGluPhe 404
QY 964 ---CCGACGACGAAATTCCTCGGTCGATTACGGCTTAAACCGCGACGCTGGGATCAGG 1020
Db 405 SerSerThrThrSerArgAsnSerIleProThrPheProGlnThrThrTyrAsnArg 424
QY 1021 GATTGGAAG 1029
Db 425 AspThrLys 427
```

RESULT 4

```
US-10-952-698-164
; Sequence 164, Application US/10952698
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natsasha
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF BLADDER CANCER, COMPOSITIONS AND METHODS
; FILE REFERENCE: FILE
; CURRENT APPLICATION NUMBER: US/10/952,698
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 164
```

```
; LENGTH: 1380
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-952-698-164

Alignment Scores:
Pred. No.: 0.0347 Length: 1380
Score: 103.50 Matches: 90
Percent Similarity: 39.01% Conservative: 59
Best Local Similarity: 23.56% Mismatches: 136
Query Match: 5.74% Indels: 97
DB: 6 Gaps: 21

US-10-089-688-1 (1-1053) x US-10-952-698-164 (1-1380)
QY 13 TTTCGGAGCTCTCAAAATCTCCAGCAGACATGTTGTAAGATCTG----- 57
Db 888 PhelysLysGluAsnGluThrLeuLysSerAspLeuAsnLeuMetGluLeuGlu 907
QY 58 -----AAGGAGAGCATGGCTTCTCGAAAGCAAGACATTTCTGATAAAAGCAGAA 111
Db 908 AlaGluLysGluArgAsnAsnLysLeuSerLeuGlnPheGluGluAspLys-----Glu 925
QY 112 AAGCTCAGAGAAGTTTCCAAAATCTGGTTGCCATGAAGAAATTTCTGTATGGCACA 171
Db 926 AsnSerSerLysGluLeuLysValLeuLysValLeuGluAlaValArgGln----- 940
QY 172 AATGAAAAAGAGCTCAGACAGACAGTAGTACTCAA-----CTTGCTCAGAACTC 222
Db 941 ---GluLysGlnLysGluThrAlaLysCysGluGlnMetAlaLysValGlnLysLeu 959
QY 223 TATAATAGTGGCTCCTTAGCACC-----CTGTAGCTGATTTACAGCTCATTGACTTT 276
Db 960 GluLysSer---LeuLeuAlaThrGluLysValLysSerSerLeuGluLysSerArgAsp 978
QY 277 GAGGCGAAAAAGACGGCTCAAAATTTCAACAATATT-----CTCAGAAGCAAAAT 330
Db 979 SerAspLysLysValValAlaAspLeuMetAsnGlnLeuGlnLeuArgThrSerVal 998
QY 331 GGTACGAGAACTCTACTGTTGAATACATCTGCACCAACAGATATTTGTCATCTTA 390
Db 999 CysGluLysThrGluThrIleAsp-----ThrLeu 1008
QY 391 TTGAAGAGGTATGAATCTCCAGAAATAGCTCTAAATTTGGAATAATTAAGAGAAATGC 450
Db 1009 LysGlnGluLeuLysAspIleAsnCysLysTyrrAsnSerAlaLeuValAspArgGluGlu 1028
QY 451 ATCAGACATGAACCACTTGCAAAAATCATTTTGGTCGGAAACAGTTTATGATTTCTTC 510
Db 1029 SerArg-----ValLeuIleLysLysGlnGluValAspIleLeu 1041
QY 511 AGATATGCGAAATGTCACATTTGACATGCTTCAGATGCTATGTCACATTCACAGAT 570
Db 1042 AspLeuLysGluThrLeuArgLeuArgIleLeuSerGluAspIleGlu-----ArgAsp 1059
QY 571 TTACTT-----ACAAGACATAAATGCTCAGTCGACAGAAATTTTG 609
Db 1060 MetLeuCysGluAspLeuAlaHisAlaThrGluGlnLeuAsnMetLeuThrGluAlaSer 1079
QY 610 GAACAGCATTTATGATGATTTTTCAGTGAATATATGAGAAGTTTACTTCATTTCAGAAAATAT 669
Db 1080 LysLysHisSerGlyLeuLeuGlnSerAlaGlnGluGluLeu----- 1093
QY 670 GTGACAAAAGACAGCTACTAGCTTCTCGTGAACCTACTACTAGTAGACACAACTTC 729
Db 1094 ---ThrLysLysGluAla-----LeuIleGlnGluLeu-----GlnHis----- 1105
QY 730 ACAATTATGACAAAATATACATCAGTAAACCTCAGAACCTCAAATTAATGATCAACTGCTG 789
Db 1106 -----LysLeuAsnGlnLysLysGluGluValGluGlnLysLysAsnGluTyr 1121
QY 790 CGAGACAAAAGTCGCAACATCCAGTTTGTAGGCCCTTTCACGTTTAAAGGTGTTGTAGCC 849
Db 1121 -----LysLeuAsnSerGlnHisTyrrGlnLysAsnIleAsnSerLeuGlnGluLeuGln 142

1122 AsnPhelysMetArgGlnLeuGlu-----HisValMetAspSerAlaAlaGlu 1137
QY 850 AATCCTTAAC-----AAGACGCGAGCCC-----ATCCTAGACATCTCTCTCAAG 891
Db 1138 AspProGlnSerProLysThrProHisPheGlnThrHisLeuAlaLysLeuGlu 1157
QY 892 AACGAGCCAAACTCATAGAG-----TTCCTC 918
Db 1158 ThrGlnGluGlnGluIleGluAspGlyArgAlaSerLysThrSerLeuGluHisLeuVal 1177
QY 919 AGCAAGTTTTCAGAACACGACGACGCGGATTGTATGACGAGTCCCTACCGACGACGATTC 978
Db 1178 ThrLysLeuAsnGluAspArg-----GluValLysAsnAla 1189
QY 979 CGGGTCGATTTACGCTTAAACCGCGACGCGTGGGATCAGGATTTGAAGACAGCAGCT 1038
Db 1190 GluIle---LeuArgMetLysGluGlnLeuArgGluMetGluAsnLeuArgLeuGluSer 1208
QY 1039 CAGCAA 1044
Db 1209 GlnGln 1210

RESULT 5
US-10-967-702-245
; Sequence 245, Application US/10967702
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Pauloski, Nicole
; APPLICANT: Taylor, Ian
; APPLICANT: Bigwood, Douglas
; FILE REFERENCE: 5176
; TITLE OF INVENTION: GENE EXPRESSION PROFILES AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/10/967,702
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/508,355
; PRIOR FILING DATE: 2003-10-03
; NUMBER OF SEQ ID NOS: 400
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 245
; LENGTH: 1583
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-967-702-245

Alignment Scores:
Pred. No.: 0.0359 Length: 1583
Score: 103.50 Matches: 84
Percent Similarity: 36.57% Conservative: 59
Best Local Similarity: 21.48% Mismatches: 133
Query Match: 5.74% Indels: 115
DB: 6 Gaps: 19

US-10-089-688-1 (1-1053) x US-10-967-702-245 (1-1583)
QY 19 AAGTCTCACAATCTCCAGCAGACATTTGTAAGAAATCTGAAGAGACATCGCTGTT--- 75
Db 43 LysTyrrSerGluAspLysAlaAsnLeuGlnLysGlnLeuGluAlaMetAsnThrGln 62
QY 76 -----CTGGAAGCAACACATTTCTGATAAAAGCAGAAAGGCT 117
Db 63 LeuGluLeuSerGluGlnLeuLysPheGlnAsnSerGluAspAsnValLysLysLeu 82
QY 118 ACAGAAGAAGTTTCCAAAATCTGTTGCCATGAAGAA---ATTCTGTAT----- 165
Db 83 GlnGluGluIleGluLysIleArgProGlyPheGluGluGlnIleLeuTyrrLeuGlnLys 102
QY 166 -----GGCAAAATGAAAAAGAG-----CCTCAGACAGACGAGTAGCT 204
Db 103 GlnLeuAspAlaThrThrAspGluLysLysGluThrValThrGlnLeuGlnAsnIleIle 122
QY 205 CAACTTCTCAGAACTCTATATAGTGGCTCTTAGCACCCTGTTAGCTGATTTACAG 264
Db 123 GluAlaAsnSerGlnHisTyrrGlnLysAsnIleAsnSerLeuGlnGluLeuGln 142
```

```
QY 265 CTC-----ATTGACTTTGAGGC 282
Db 143 LeuylsAlaIleHisGlnGluValylsGluLeuMetCysGlnIleGluAlaSerAla 162
QY 283 AAAAAAGACGTGGCTCAAAATT-----TTCAACAATATTCTCAGA 321
Db 163 LysGluHisGluAlaGluIleAsnLysLeuAsnGluLeuLysGluAsnLeuValylsGln 182
QY 322 AGACAAATTGTCAGAGACTTCCTACTGTTGAATATCATCTGCACCAACAGAAATATTG 381
Db 183 CysGluAlaSerGluLysAsnIleGlnLysAlsYrGluCysGluLeuGluAsnLeuArg 202
QY 382 TTCATGTTATTGAAGGTATGAATCTCCAGAAATAGCTCTAAATTTGGNATATCTTA 441
Db 203 LysAlaThrSerAsnAlaAsnGlnAspAsnGlnIle-----CysSerIleLeuLeu 219
QY 442 AGAATATGCATCAGACATGAACCTTGCAGAAATCATTTTGTGTCGGAACAGTTTAT 501
Db 220 GlnGlu----- 221
QY 502 GATTCTTCAGATATGCGAAATGTCACATTTGCAGATAGCTTCAGATGCAATTTGCCACA 561
Db 222 -----AsnThrPheValGluGlnValValAsnGluLysValylsHisLeuGluAspThr 239
QY 562 TTCAGAGTTTACTTACAGACATAAATTGCTAGTCAGTCAGAAATTT----- 606
Db 240 LeuLysGluLeuGluSerGlnHisSerIleLeuLysAspGluValThrTyrMetAsnAsn 259
QY 607 -----TTGGA-----CAGCATTAT--GATAGATTTTTCAGTGAATAT 642
Db 260 LeuLysLeuLysLeuGluMetAspAlaGlnHisIleLysAspGluPhePheHisGluArg 279
QY 643 GAGAGTTACTTCTCAT-----TCAGAAATTTATGTGCAAAAAGACAGTCA----- 687
Db 280 GluAspLeuGluPheLysIleAsnGluLeuLeuLeuAlaLysGluGlnGlnCysVal 299
QY 688 -----CTGAGCTTCTCGGTGAACATA-----CTACTAGAT 717
Db 300 IleGluLysLeuLysSerGluLeuAlaGlyLeuAsnLysGlnPheCysTyrThrValGlu 319
QY 718 AGACACAATCTCAAAATTTATGACA-----AAATACATC----- 750
Db 320 GlnHisAsnArgGluValGlnSerLeuLysGluGlnHisGlnLysGluIleSerGluLeu 339
QY 751 -----AGTAAACCTGAGAACCTCAAAATTAATGATGAACCTG--- 786
Db 340 AsnGluThrPheLeuSerAspSerGluLysGluLysLeuThrLeuMetPheGluIleGln 359
QY 787 ---CTGCGACACAAAAGTCGACATCCAGTTTGAGGCCCTTTCACGTTTTTAAGGTGTTT 843
Db 360 GlyLeuLysGluGlnCysGluAsnLeuGlnGlnLys-----GlnGluAla 375
QY 844 GTAGCCAACTTAACAGACGCGACCCCTCTAGACATCTCTCAAGAACCGGCCAAA 903
Db 376 IleLeuAsnTyrGluSerLeuArgGluIleMetGluLeu-----GlnThrGlu 392
QY 904 CTCATAGAGTTCTCAGCAAGTTTTCAGAACGAC 936
Db 393 LeuGlyGluSerAlaGlyLysIleSerGlnGlu 403
```

RESULT 6

```
US-10-732-923-8275
; Sequence 8275, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES.
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
```

```
; SEQ ID NO 8275
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Mycoplasma genitalium
US-10-732-923-8275

Alignment Scores:
Pred. No.: 0.0353 Length: 600
Score: 102.50 Matches: 73
Percent Similarity: 38.80% Conservative: 69
Best Local Similarity: 19.95% Mismatches: 151
Query Match: 5.68% Indels: 73
Db: 6 Gaps: 14
```

US-10-089-688-1 (1-1053) x US-10-732-923-8275 (1-600)

```
QY 88 GACATTCTCATATAAAAGACAGAA----- 111
Db 66 AspLeuAsnAspGluAsnAlaLysAlaLeuIleSerPheTrpSerGlnIleAspLeuArg 85
QY 112 AAGGTACAGAGAAGTTTCCAAAAATCTGGTTGCCATGAAAGAAATTTCTGTATGCG--- 168
Db 86 AsnAlaThrCysAspAlaAsnAsnLeuLeuAlaLysTyrHisLeuPheGluVal 105
QY 169 -----ACAAATGMAAAAGAGCCTCAGACAGAAGCA 198
Db 106 ValAspThrIleSerLysTyrThrAspValThrAsnLeuLeuAspGluLysThrAlaLeu 125
QY 199 GTAGCTCAACTCTCAAGAA---CTCTATAATAGTGGGCTCTTAGCACCTGGTAGCT 255
Db 126 GlnLeuLysIleAlaSerGluAsnGlnPheHisSerLeuAspAsnAsnGlnIleAsn 145
QY 256 GATTACAGCTCAATTGACTTTGAGGGCAAAAAGACGTGGCTCAAAATT----- 303
Db 146 AsnLeuArgLysTyrGlnLysGluValLysTrpIleArgAlaLeuGluAspAsnGln 165
QY 304 TTCACAAATATTTCTCAGAGACAAATTTGTACAGAACTCTACTGTGTAATACATCTGC 363
Db 166 PheGlyGlyIleLeuAlaAspGluMetGly-----LeuGly 177
QY 364 ACCCAACAGAAATATTTGTTTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTCTA 423
Db 178 LysThrAlaGlnValIlePheAlaMetLeuAspSerTyrGlnSerThrLysSerLeuLeu 197
QY 424 AATTGTGGAATATGTTAAAGAGAATGCATCAGACATGAACCACTTGCAGAAATCATTTTG 483
Db 198 ProSerLeuIleIleVal-----ProAlaSerLeuLeuLeuAsn 210
QY 484 TGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACATTTGACATAGCT 543
Db 211 TrpLysSerGluPheGlnLysPheAlaProHisValLysIleValThr----- 226
QY 544 TCAGATGTCATTGCCACATTCAAGGATTTTACTTACAAGACATAAATGCTCAGTCGAGAA 603
Db 227 -----AlaAsnGlyAsnPheLysGluArgSerGlnValTyrGluSerLeuLysAsnGln 244
QY 604 TTTTGTGAACAGCATATGATGATTTTTCAGTGAATAT-----GAGAAAGTTA 651
Db 245 IleLeuLeuMetSerPheAsnValLeuArgSerAspIleLysTrpIleSerGlnLysLys 264
QY 652 CTTTCAAT-----TCAGAAATATTGTGACAAAA 678
Db 265 PheHisTyrValValIleAspGluAlaGlnGlyIleLysAsnGluAsnSerThrValThr 284
QY 679 AGACACTGCTGAAGCTTCTCGGTGAACCTACTACTAGATAGACACAACTTCACAATATG 738
Db 285 LysAlaAlaLysLysIleLysGlyAsnPheCysLeuAlaLeuThrGlyThrProIleGlu 304
QY 739 ACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG---AACCTGCTGCGAGAC 795
Db 305 AsnArgLeuLeuAspLeuTrpSerCysPheAspPheValLeuProAsnPheLeuGlyAsn 324
QY 796 AAAAGTCGC---AACATCCAGTTTGGAGCCTTTACGTTTAAAGGTGTTGTAGCCAAT 852
```

```
Db 325 LysLeuSerPheSerGlnPheGluLysGluLysAsnAspGluSerPheGlnLysLeu 344
QY 853 CCTAAAGAGAGCGCCATCTCTAGACATCTCTCTCAAGAACACGCGCCAAACTCATAGAG 912
Db 345 MetLysLysThrSerProPhe-----IleLeuArgArgThrLysAsnLysValLeuLys 362
QY 913 TTCTCTAGCAAGTTTCAAGACGACGAGCGAGTGTATGACGAGTTCGCTACCGACGACG 972
Db 363 GluLeuProLys-----LysIleIleThrAspIleTyr-----ValGluLeuSer 377
QY 973 AATTCGGGTCGATTACCGGTTAAACCGCGACGCGTGGATCAGGGATTTGAAGAGA 1032
Db 378 GluGluHisGlnLysLeuTyrAspLysGlnLysThrAspGlyLeuLysGluLysGlu 397
QY 1033 CCAGCTCAGCAAGAAGCT 1050
Db 398 SerAspAlaLysAsnAla 403
Db 398 SerAspAlaLysAsnAla 403
RESULT 7
US-10-732-923-3305
; Sequence 3305, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3305
; LENGTH: 921
; TYPE: PRT
; ORGANISM: Fusobacterium nucleatum subsp. vincentii ATCC 49256
; US-10-732-923-3305
Alignment Scores:
Pred. No.: 0.0439 Length: 921
Score: 102.00 Matches: 79
Percent Similarity: 33.93% Conservative: 53
Best Local Similarity: 20.31% Mismatches: 137
Query Match: 5.66% Indels: 120
DB: 6 Gaps: 16
US-10-089-688-1 (1-1053) x US-10-732-923-3305 (1-921)
QY 13 TTGGG---AAGTCTCACAATCTCCAGCAGACATTTGTGAAGAAT----- 54
Db 65 PheGlyGluLysSerSerLysValAspPheIleAlaAsnAspGlyArgGluTyr 84
QY 55 -----CTGAAGCAGGACATG 69
Db 85 AsnLeuLysThrGluPhePheLysThrLysProLysLysGlnThrLeuLysAspMetIle 104
QY 70 GCTGTTCTGGAAGAACAGACATTTCTCATATAAAGCAGAA-----AAG 114
Db 105 GlySerGluTyrAspGlyAspIleGlnLysLeuGluLysGlyIleLysLys 124
QY 115 GCTCAGAGAAGTTTCAAAATCTGTTGTCATGAAA-----GAAATCTGTATGGC 168
Db 125 GlyPheGluGluThrTyrGluAsnIleValIleAlaLysGlnAsnGluPheIleAsnIle 144
QY 169 ACAATGAAAAGAGCCTCAGACAGACAGTA-----GCTCAACTT 210
Db 145 PheLysAlaLysProLysAspArgGluGluIlePheAsnLysIlePheAsnThrGlnIle 164
QY 211 GCTCAAGACTCTATAATAGTGGGTC----- 237
Db 165 TyrLysGluMetTyrAspSerPheLysGluAlaIleAspLysTyrLysSerGluLys 184
QY 238 -----CTTAGCACCTGTAGTCAATTTACAGTCACTTGTAGGGGCAAAAAGAC 291
Db 238 -----CTTAGCACCTGTAGTCAATTTACAGTCACTTGTAGGGGCAAAAAGAC 291
```

```
Db 185 GluAsnLeuAspSerLysIleAsnSerLeuLys-----GluAsnMetGluLys 200
QY 292 GTGGCTCAAAATTTCAACAATATT-----CTCAGAGACAAATTTGGTAGC 336
Db 201 LysGluGlnIleThrAsnPheLeuLysGluLysGluValGluLysAsnLeuGlnAsp 220
QY 337 AGAATCTCTACTGTTGAATACATCTGCAACCAAGAAATATTTGTTCAATTTATTTGAAA 396
Db 221 ArgPheLysAsnIleAsnValSerLysAsn-----LeuGluAsnGluLysLys 237
QY 397 GGGTGAATCTCCAGAAATAGCTCTAAATTTGGAATA----- 435
Db 238 AspTyrGluThrGluIleGluAsnAsnLeuIleLysAsnIleLysAspGluGlu 257
QY 436 -----ATGTTAAGAGATGATCATGATGCAATTCAGGATTTACTTACAAGACAT 471
Db 258 AsnLysIleLysLysTyrLeuAsnIleLysGluAsnIleIleGluAlaLysGlnAla 277
QY 472 AAA-----ATCATTTTGTGGTGGAAACAGCATTTATGATAGATTTTTCAGTGAATAG 525
Db 278 LysLysSerLysIleIleValLysGluThrGluLysSerTyrPheGluTyrLeuGluIle 297
QY 526 TCACATTTGACATAGCTTCAGATGCAATTCGACATTCAGGATTTACTTACAAGACAT 585
Db 298 GluAsn-----ArgLeuLysAsp----- 303
QY 586 AAATTGCTCAGTCAGCAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATAG 645
Db 304 -----LeuArgGluAsnLeuAspAsnLeuLysGlnLys 316
QY 646 AAGTTACTTCAATTCAGAAAATTTATGTCACAAAGACAGTCACTGAGCTTCTCGGTGAA 705
Db 317 LeuAsnIleGlnTyrGlnAsnAsnIleGluLys----- 327
QY 706 CTACTACTAGATAGACACACTTCACAAATTTATGACAAATATACATCAGTAAACCTGAGAA 765
Db 328 LeuGluLeuSerAsnLysAsnLeuLysVal-----AspIleSerAsn 341
QY 766 CTCAAAATTAATGATGAACCTGCTGCGAGACAAAGTCGCAACATCCAGTTTGGGCTTT 825
Db 342 LeuGluGluAsnIleSerLysAsnSerGluLysLysGluAsnLeuGluSerGlu----- 359
QY 826 CACGTTTTTAAGGTGTTGTAGCAATCTTAACAGACGCGCCATCTCTAGACATCCTC 885
Db 360 -----IleSerAsnLeuLysIleLysGluLysLeuAspLeuAspLeuLys 373
QY 886 CTCAAGAACACGCGCCAACTCATAGATTTCTCAGCAAGTTTCAGAACGACGACGCGAT 945
Db 374 LeuLysLysTyrThrThrLeuLeuAspGluLeuGluLysLysIleLysLysGlu 393
QY 946 TGTATGACGATTCCTGCTACCGACGACG 972
Db 394 LysSerAspLysLysLeuLysThrThr 402
RESULT 8
US-10-886-744-1
; Sequence 1, Application US/10886744
; GENERAL INFORMATION:
; APPLICANT: Guan, Jun-Lin
; APPLICANT: Abbi, Smita
; TITLE OF INVENTION: METHODS OF REGULATING FOCAL ADHESION KINASE AND ITS
; TITLE OF INVENTION: ASSOCIATED CELLULAR FUNCTIONS BY FAK FAMILY-INTERACTING
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 19603/4451
; CURRENT APPLICATION NUMBER: US/10/886,744
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/486,159
; PRIOR FILING DATE: 2003-07-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1591
; TYPE: PRT
```

; ORGANISM: Homo sapiens					
US-10-886-744-1					
Alignment Scores:					
Pred. No.:	0.0503	Length:	1591		
Score:	102.00	Matches:	76		
Percent Similarity:	36.36%	Conservative:	60		
Best Local Similarity:	20.32%	Mismatches:	114		
Query Match:	5.66%	Gaps:	125		
DB:	6	Gaps:	16		
US-10-089-688-1 (1-1053) x US-10-886-744-1 (1-1591)					
QY	25	CACAAATCTCCAGCAGACATTTGCAAGAATCTGAAGCAGACATCGCTTCTGTGGAAGAAG	84		
Db	1162	HISAsnGlnAlaPheGluIleGluLeuAAsnLeuLysGluGlnIleIleGluLeuGlnSer	1181		
QY	85	CAGACATTTCGTAT-----AAAAAGCAGAAAAGGCTACAGAGAA	126		
Db	1182	LysLeuAspSerGluLeuSerAlaLeuGluArgGlnLysAspGluLysIleThrGlnGln	1201		
QY	127	GT TTC AAAAATCTGGTTGCCATGAAGAATACTGTATGCACAATAGAAAAGAGCCCT	186		
Db	1202	-----GluGlu	1203		
QY	187	CAGACAGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAAGTGGCTCTTAGCACCC	246		
Db	1204	LysrtyrGluAlaIleIleGln-----Asn	1211		
QY	247	CTGGTAGCTCATTTATCAGCTCAITGACTTTGAGGGCAAAGACGTGGCTCAAAATTTTC	306		
Db	1212	LeuGluLysAspArgGlnLysLeuValSerSerGlnGlnGlnAspArgGluGlnLeuIle	1231		
QY	307	AACAATATCTCAGAGACAAATGGTACGAGAACTCTACTGTGTAATACATCTGCACC	366		
Db	1232	GlnLysLeu-----AsnCysGlu	1237		
QY	367	CAACAGAAATATTTTGTTCATGTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAAT	426		
Db	1238	LysAspGluAlaIleGlnThrAlaLeuLysGluPhelys-----	1250		
QY	427	TGTGGAATATGTTTAAGAGAGAAATCAGACATCAGAACCACTTGCAAAAATCAITTTGTGG	486		
Db	1251	-----LeuGluArgGluValGluLysGluLeuLeuGluLysVal-----	1264		
QY	487	TCGGAACAGTTTTATGATTTCTTCAGATATGTCGAATG-----	525		
Db	1265	-----LysHisLeuGluAsnGlnIleAlaLysSerProAla	1276		
QY	526	-----TCAACATTTTGACATAGCTTCAGATGCATTTGCCACATTCAGAGGATTTACTTACA	579		
Db	1277	IleAspSerThrArgGlyAspSerSerLeuValAlaGluLeuGlnGluLysLeuGln	1296		
QY	580	AGACATAAATPGCTCAGTCAGAGAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAA	639		
Db	1297	GluGluLys-----AlaLysPheLeuGluGln-----LeuGluGlu	1308		
QY	640	TATGAGAAGTTACTTCATTCAGAAAATATTGTGACAAAAAGACAGTCACCTGAAGCTTCTC	699		
Db	1309	GlnGluLys-----ArgLysAsnGluGluMetGlnAsnValArgThrSerLeuIle	1325		
QY	700	GGTGAATCTACTAGATAGACACAACCTTC--ACAATTATGACAAAATACATCAGTAAA	756		
Db	1326	AlaGlu-----GlnGlnThrAsnPheAsnThrValLeuThrArgGluLysMetArg	1342		
QY	757	CTGAGAAGCTCAAAATTAATGATGAACCTGCTCGCAGACAAAAAGTCGCAACCTCAGTTT	816		
Db	1343	LysGluAsnIle-----IleAsnAspLeuSerAspLys-----	1353		
QY	817	GAGGCCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCTTAACAGACGCGAGCCCATCCTA	876		
Db	1354	-----LeuLysSerThrMetGlnGlnGlnGluArgAspLysAspLeuIle	1368		

Db 1238 LysAspGluAlaIleGlnThrAlaLeuLysGluPheLys----- 1250
QY 427 TGTGGAATATGTAAGAGAAATGCATCAGACATGAACCACTTCCAAAAATCATTTGTGG 486
Db 1251 -----LeuGluArgGluValValGluLysGluLeuLysVal----- 1264
QY 487 TCGGAACAGTTTATGATTTCTTCAGATATGTCGAATG----- 525
Db 1265 -----LysHisLeuGluAsnGlnIleAlaLysSerProAla 1276
QY 526 -----TCACATTTGACATAGCTTCAGATGCAATTTGCCACATCAAGGATTTACTTACA 579
Db 1277 IleAspSerThrArgGlyAspSerSerLeuValAlaGluLeuGlnLysLeuGln 1296
QY 580 AGACATAAATGCTCAGTGCAGAAATTTTGGACACAGCATATGATAGATTTTTCAGTGAA 639
Db 1297 GluGluLys-----AlaLysPheLeuGluGln-----LeuGluGlu 1308
QY 640 TATGAGAGTTACTTCTTATTCAGAAATATGTCACAAAAGACAGTCACCTGAAGCTTCTC 699
Db 1309 GlnGluLys-----ArgLysAsnGluGluMetGlnAsnValArgThrSerLeuIle 1325
QY 700 GTGAACACTACTAGATAGACACAACTTC---ACAATTATGACAAATACATCAGTAA 756
Db 1326 AlaGlu-----GlnGlnThrAsnPheAsnThrValLeuThrArgGluLysMetArg 1342
QY 757 CTTGAGAACCTCAAAATTAATGATGAACTGCTGCGAGACAAAAGTCGCAACATCCAGTTT 816
Db 1343 LysGluAsnIle-----IleAsnAspLysSerAspLys----- 1353
QY 817 GAGCCCTTTCACGTTTAAAGTGTGTTAGTCCATCTTAACAGACGACGCCCATCTTA 876
Db 1354 -----LeuLysSerThrMetGlnGlnGluArgAspLysAspLeuIle 1368
QY 877 GACATCTCTCAAGAACAGCCAACTCATAGATGTTCTTCAGCAAGTTTTCAGAACGAC 936
Db 1369 GluSerLeuSerGluAspArgAlaArgLeuLeuGluGluLysLysLysLeuGluGlu 1388
QY 937 AGGACGATGTTATGAGCAGTTCC---GTACCGACGAC----- 971
Db 1389 ValSerLysLeuArgSerSerPheValProSerProtyrValAlaThrAlaProGlu 1408
QY 972 -----GAATCCCGGGTTCGATTACGCGTTTAAACCGCGGACGCG 1010
Db 1409 LeuTyrglyAlaCysAlaProGluLeuProGluProGlyGluSerAspArgSerAlaValGluThr 1428
QY 1011 TGGGATCAGGATTTGAAGACACCGCTCAGCAAGAAGCTTA 1052
Db 1429 AlaAspGluGlyArgValAspSerAlaMetGluThrSerMet 1442

RESULT 10

US-10-732-923-10741
; Sequence 10741, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10741
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Entamoeba histolytica
US-10-732-923-10741

Alignment Scores:

Pred. No.: 0.0628 Length: 409
Score: 99.50 Matches: 66
Percent Similarity: 33.12% Conservative: 37

Best Local Similarity: 21.22% Mismatches: 111
Query Match: 5.52% Indels: 97
DB: 6 Gaps: 13
US-10-089-688-1 (1-1053) x US-10-732-923-10741 (1-409)
QY 16 GGAAGTCTCAAAATCTCCAGCAGACATTTGAGAAATCTGAAGGAGCAGCATGGCTGTT 75
Db 2 GlyAsnSerAsnHisLysProGluLysLysAsnAsnLysGlyLysGlyLysGlyLys 21
QY 76 CTGGAACCAAGACATTTCTGATATAAAAGCAGAAAAG---GCTACAGAAAGAGTTTCC 132
Db 22 ArgLysAsnGluLysLysAspGluThrAsnGluLysGluAlaIleGluProThr 41
QY 133 AAAATCTGTTGCCATGAAGAAATTTCTATGGCACAATAATCAAAAAGAGCTCAGACA 192
Db 42 ProAsn-----GluGluLysGluProCysThr 50
QY 193 GAAGCAGTAGCT-----CACTTGCTCAA----- 216
Db 51 AspThrLysProAsnGluAspAsnGlnIleValLysGlyLysGluGlyLysGluGlu 70
QY 217 -----CAACTCTATAATAGTGGCTCCTTAGCACCCCTGCTAGCTGATTTACAGCTCAT 270
Db 71 TyrlleGluLeuTyrrAsnGluGlyLysLysValThrGlnAspAspPheGluLeu 90
QY 271 GACTTTGAGGCG-----AAAAAGACGTGGCT 297
Db 91 LysValIleGlyArgGlySerPheGlyLysValMetMetValLysLysLysAspGly 110
QY 298 CAAATTTTC---AACAAATATCTCAGAGACAAATTTGGTACGAGAACTCTACTGTTGAA 354
Db 111 ArgIlePheAlaMetLysIleLeuArgLysAspIleValLysGluArgLysGlnValAsp 130
QY 355 TACATCTGCACCAACAGAAATATTTGTCATGTTTGAAGGGTATGAATCTCCAGAA 414
Db 131 HisThrLysAlaGluLysAsnValLeuMetGlnLeu----- 142
QY 415 ATAGCTCTAAATTTGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACACCATTTG 474
Db 143 -----HisHisProPheIleVal 148
QY 475 ATCATTTTGTGTCGGAACAGTTTATGATTTTCTTCAGATATGTCGAAATGTCAACATTT 534
Db 149 LysLeuTyrrAlaPheGln----- 155
QY 535 GACATAGCTTCAGATGATTTGCCACATTCAGGATTTACTTACAGACATAAATGCTC 594
Db 156 -----ThrThrAspLysLeuTyrrMetValMetAspPheValAsnGlyGlyLeu 172
QY 595 AGTCAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAGTTACTT 654
Db 173 -----PheTyrrHisLeuLysAsnGluAsnCysPheSerGluGluArgAlaLysPhe 189
QY 655 CATTCAAGAAATATGTGACAAAAGACAGCTCACTGAAGCTTCTCGGTGAACTACTACTA 714
Db 190 TyrAlaAlaGluIleAlaThr-----ValLeuIle 199
QY 715 GATAGACACACTTCACAAATTTATGACAAAATATACATAGTAACTGAGAACTCAAAATTA 774
Db 200 HisIleHisSerLeuGlyIleIleTyrrArgAspLeu---LysProGlu----- 214
QY 775 ATGATGAACCTGCTGCGAGACAAAGTCGCAACATCCAGTTTGTAGGCGCTTTCAGTTT 834
Db 215 -----AsnIleLeuLeuAspAsnThrGlyAsnIleValIleThrAspPheGlyLeuSer 232
QY 835 AAGGTGTTGTAGCCAATCTTAAACAGCCGAG 867
Db 233 LysGlnLeuAlaAlaGly---GluGluThrGln 242

RESULT 11

US-10-732-923-8676
; Sequence 8676, Application US/10732923

```

; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8676
; LENGTH: 1251
; TYPE: PRT
; ORGANISM: Encephalitozoon cuniculi
US-10-732-923-8676

Alignment Scores:
Pred. No.: 0.0926 Length: 1251
Score: 99.00 Matches: 81
Percent Similarity: 36.36% Conservative: 59
Best Local Similarity: 21.04% Mismatches: 127
Query Match: 5.49% Indels: 118
DB: 6 Gaps: 20

US-10-089-688-1 (1-1053) x US-10-732-923-8676 (1-1251)

QY 1 ATGCCGTTCCGTTGGAGTCTCACAATCT---CCAGCAGACATGTGAAGATCTG 57
D 800 MetThrPheGluGluGlnLysSerGluLysThrArgLysLysAspAlaIleGluThrLeu 819
QY 58 AAGGAGACATGGCT-----GTTCTGAAAGCAAGACATTTCTGTAAAGCA 108
D 820 LeuArgLysGlyAlaTyrGlyValLeuMetGluThrAspAlaSerSerArgLysPhe 839
QY 109 GAAAGGCTACAGAAAGATTTCGATGCTGCTGCAAGAAATCTCTATGGC 168
D 840 -----CysGluGluAspIleAspGlnIleGluLysArgThrLysIleValLysHis 857
QY 169 ACAATGAAAGAGCTTCAGAGACAGTACTCACTTGTCTCAAGAACTTATAAT 228
D 858 SerAspGlyGlyAsnValPheSerLysAlaSerPheGlnValGluGluGluIleAspAsp 877
QY 229 AGTGGGCTCTACGACCCCTGGTAGCTGATTACAGCTCATTGACTTTGAGGCAAA--- 285
D 878 ProAspPheTrpGluAsnLeuLeuSerLysArgLysSerGluGluSerGluGlyArgIle 897
QY 286 -----AAAGACGTGGCTCAA 300
D 898 ArgArgGlnMetArgArgLeuAlaArgGluGlyLeuSerSerGluAspIleSerGlu 917
QY 301 ATTTTCAACAATATTCAGAGACAAATGGTAGAGAACTCTCTACTGTGTGAATACATC 360
D 918 Ile---AspGluLeuLeuAsnThrLysLeuGluSerGluAspGluAsnValPheTyr--- 935
QY 361 TGCACCCACAGATATTTTGTTCATGTTA-----TTGAAA 396
D 936 ---GluAsnGlnCysLeuLeuIlePheLeuGlyValLeuArgPheGlyProArgSerVal 954
QY 397 GGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGGAATAATGTTAAGAGAATGCATCAGA 456
D 955 GlyLeuGluIleProSerLeuThr----- 962
QY 457 CATGAACCACTT-----GCAAAAATCATTTTGTGGTGGACAGTTT 498
D 963 HisGluAspLeuLysAspGlyValGluLysAlaHisIleLeuProGlnSerProArgPhe 982
QY 499 TATGAT-----TTCTTCAGATATGTCGAATATGCAATTTGAC----- 537
D 983 LysAspThrAspAlaTyrPheArgTyrLeuValLysTyrThrValAspGlnPheGlnSer 1002
QY 538 -----ATAGCTTCAGATGCAATTTGCCACATTCAGGATTTACTTACAGACATAATTTG 591
D 1003 ThrLysValArgSerAsp-----PheSerAspGlyIle----- 1013

; SEQUENCE 3307, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3307
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Methanocaldococcus jannaschii
US-10-732-923-3307

Alignment Scores:
Pred. No.: 0.11 Length: 1005
Score: 98.00 Matches: 84
Percent Similarity: 42.99% Conservative: 57
Best Local Similarity: 25.61% Mismatches: 120
Query Match: 5.44% Indels: 67
DB: 6 Gaps: 18

US-10-089-688-1 (1-1053) x US-10-732-923-3307 (1-1005)

QY 49 AAGATCTGAAGAGAGCATGCTGTCTGGAAGACAA-----GACATT 93
D 194 LysGluLeuLysAsnLysMetSerGlnLeuGluLysAsnLysLysLeuMetGluIle 213
QY 94 TCTGTATAA-----AAAGCAGAAAGGCTACAGAGAACTTTCCAAA-----AATCTG 141
D 214 AsnAspLysLeuAsnLysIleLysLysGluPheGluAspIleGluLysLeuPheAsnGlu 233
QY 142 GTTGCCATGAAGAAATTTCTGTATGGC-----ACAATGAAAGAGCGCTCAGACAGAA 195
D 234 TrpGluAsnLysLysLeuLeuTyrGluLysPheIleAsnLysLeuGluLysArg 253
QY 196 GCATAGCTCAACTTGCTCAAGAACTC-----TATAATAGTGGGCTCTTAGC 243
```

```
Db 254 AlaleuGluLeuLysAsnGlnGluLeuLysIleLeuGluTyrAsp-----LeuAsn 270
Qy 244 ACCGTGTAGCTGATTACAGCTCATTGACCTTTGAGGGGCAAAAGACGGTGGCTCAAAAT 303
Db 271 ThrValValGluAlaArgGlu-ThrLeuAsn--ArgHisLysAspGluTyrGluLysTy 289
Qy 304 TTCACAAATATTCAGAGAACAAATTTGGTACGAGAACTCTACTGTTGAATACATCTGC 363
Db 289 rLysSerLeuValAspGluIleArgLysIleGluSerArgLeu----- 303
Qy 364 ACCCAACAGAAATATTTGTTGATCTATTGAAAGGTATGAATCTCCAGAAATAGCTCTA 423
Db 304 -----ArgGluLeuLysSerHisTyrGluAspTyrLeuLysLeuThrLysGlnLeuG 321
Qy 424 AATTGTGGAATAATGTTAAGAGAAATGCATCAGACATCAACACCTTGCAAAATCATTTTG 483
Db 321 uileileLys-----GlyAspIleGluLysLeu--LysGluPheI 334
Qy 484 TGGTCGGAACAGTTTATGATTTCTTCAGAT-----ATGTCGAATGTCA 528
Db 334 eAsnLysSerLysTyrArgAspAspIleAspAsnLeuAspThrLeuLeuAsnLysIleLy 354
Qy 529 ACATTTCACATAGCTTCAGATGCTATTTGCCACATTCAAGGATTTTACTTACAAGACATAA 588
Db 354 sAsp--GluIle-----GluArgValGluThrIleLysAspLeuLeuGluLys 371
Qy 589 TTGCTCAGTCAGAAATTTTGGAA---CAGCATTTATGATAGATTTTTCAGTGA-- 639
Db 372 AsnLeuAsnGluGluIleGluLysIleGluLysTyrLysArgIleCysGluGluCysLys 391
Qy 640 -----TATGAGAGTTTACTTCATTCAGAAATTTATGTGACAAAAGACATCCTCAAG 693
Db 392 GluTyrTyrGluLysTyrLeuGluLeuGluLysAlaValGluTyrAsnLysLeuThr 411
Qy 694 CTTCCTCGTGAATCTACTACTAGATAGACACAACTTCACAATTTATGACAAAATATACATCAGT 753
Db 412 LeuGluTyrIleThrLeuLeuGluGlu-----LysLysSerIleGlu 425
Qy 754 AAACCTGAGAAC---CTCAAAATTAATGATGAACCTGCTCGAGACAAAAGTCGCAACATC 810
Db 426 LysAsnIleAsnAspLeuGluThrArgIleAsnLysLeuLeuGluGluThrLysAsnIle 445
Qy 811 CAGTTTGGGCTTTCAGCTTTTAAAGTGTGTTGAGCCATCTCAACAG----- 861
Db 446 AspIleGluSer-----IleGluAsnSerLeuLysGluIleGlu 458
Qy 862 --AGCGAGCCCATCTAGACATCTCTCTCAAGAACAGGCGCAAACTCATAGAGTTCTCTC 918
Db 459 GluLysLysLysValLeuGluAsnLeuGlnLysGluLysIleGluLeuAsnLysLysLeu 478
Qy 919 AGCAAGTTTCAGAACGAC 936
Db 479 GlyGluIleAsnSerGlu 484
```

RESULT 13

```
US-10-732-923-8707
; Sequence 8707, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8707
; LENGTH: 1359
; TYPE: PRP
; ORGANISM: Caenorhabditis elegans
US-10-732-923-8707
```

```
Alignment Scores:
Pred. No.: 0.118 Length: 1359
Score: 98.00 Matches: 82
Percent Similarity: 39.25% Conservative: 64
Best Local Similarity: 22.04% Mismatches: 142
Query Match: 5.44% Indels: 84
DB: 6 Caps: 22
```

US-10-089-688-1 (1-1053) x US-10-732-923-8707 (1-1359)

```
Qy 19 AAGTCTCAAAATCTCCAGACAGACATTGTGAAGATCTGAAGAGACATGGCTGTTCTG 78
Db 365 LysLysLysGluLysProGluLysLysLysLysGlyIleIleMetAspSerSerLysLeu 384
Qy 79 GAAAACCAAGACATTTCTGATATAAAAGACGAAAGGCTACAGAGAAGTTTCCAAAAT 138
Db 385 GlnLysGluThrIleAsp-----AlaGluArgAlaGluLysGluArgLysArg 401
Qy 139 CTGTTGCCATGAAGAAATTTCTGTATGGC-----ACAAATGAAAAGAGCCTTCAGACA 192
Db 402 LeuGluLysLysGlnLysGluPheAsnGlyIleValLeuGluGluLysLeuThr 421
Qy 193 GAAGCAGTAGCTCAACTGCTCAAGAACTCTATAATAGTGGCTCTTAGCACCTCGTA 252
Db 422 GluMetLeuThrGlyThrSerSerGlnArgLysLysSerValVal-----LeuAsp 439
Qy 253 GCTGATTTACAGCTCATTGACTTTGAGGCAAAAGACGCTGCTCAAAATTTCAACAAT 312
Db 440 ProAspSerSerThrValAspGluLysLysProVal-----GluValHisAsnSer 458
Qy 313 ATTCTCAGAGACAAATTTGTTAGAGAACTCTACTGTTGAA-----TACATCTGCACC 366
Db 459 LeuValArgIleLeuLysProHisGlnAlaHisGlyIleGlnPheMetTyrAspCysAla 478
Qy 367 CAACAGAAATTTTGTTCATGTTATGAAAGGGTGAATCTCCAGAAATAGCTCTAAAT 426
Db 479 CysGluSerLeuAspArgLeuAspThrGluGly-----SerGlyGlyIleLeuAlaHis 496
Qy 427 TGT---CGAATA-----ATGTTAAGAGAATGCATCAGACAT 459
Db 497 CysMetGlyLeuGlyLysThrLeuGlnValIleThrPheLeuHisThrValLeuMetHis 516
Qy 460 GAACCACTCTGCAAAA-----ATCATTT 480
Db 517 GluLysIleGlyGluLysCysLysArgValLeuValValProLysAsnValIleIle 536
Qy 481 TTGTGTCGGAACAGTTTATGATTTCTTC-----AGATATGTCGAAATGTCACATTT 534
Db 537 AsnTrpPheLysGluPheGlnLysTrpLeuValAspAsnAspGluGluLeuAspThrIle 556
Qy 535 GACATAGCTTCA---GATGCATTTGCCACATTTCAAGGATTTACTTACAGACATAAAATTG 591
Db 557 AspValAsnGluLeuAspSerTyrLysThrIleGluAsp-----ArgArgArgAla 573
Qy 592 CTGAGTCGAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAGTTA 651
Db 574 Leu-----LysAla 576
Qy 652 CTTCATTTCAGAAAATTTATGACAAAAGACAGTCATCAGAGCTTCTCGGTGAATCTACTA 711
Db 577 TrpHisSer-----SerLysThrProSerValMetIleIleGlyTyrAspLeu 592
Qy 712 CTAGATAGACACACTTTCACAAATTTATGACAAAATACATCATGTAACCTGACACCTCAA 771
Db 593 -----PheArgIleLeuThr---ValGluAspAspProLysLysLysLys 606
Qy 772 TTAATGATGAACCTGCTCGAGACAAAAGTCCCAACATCTCCAGTTTGTAG---GCCTTTCAC 828
Db 607 -----ProLysAsnArgAsnArgArgLeuGluLysAlaLysGlu 619
Qy 829 GTTTTAAAGTGTGTTAGCCCAATCTTAAACAAGACGCGACGCCCATCTTAGACATCTCTC--- 885
Db 620 AspPheArgLysTyrLeuGlnAsnProGlyProAspMetValValCysAspGluAlaHis 639
```

```

QY 886 ---CTCAAGAAC---CAGGCCAACTCATAGATTCCTCAGCAAGTTTCAAGACGAGG 939
Db 640 LysLeuLysAsnAspSerAlaLeuSerLysCysMetValLysIleLeuThrLysArg 659
QY 940 ACGGATTGTATGACGAGTCCGTACCGACGACGAAATCCCGGGTCGATTTACGC----- 993
Db 660 ArgIleCysLeuThrGlyThr---ProLeuGlnAsnAsnLeuMetGluTyrHisCysMet 678
QY 994 -----GTTAAACCGCGACGCGTGGGATCAGG 1020
Db 679 ValAsnPheValLysProGlyLeuLeuGlyThrLys 690

RESULT 14
US-10-732-923-8708
; Sequence 8708, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8708
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-8708

Alignment Scores:
Pred. No.: 0.118 Length: 1359
Score: 98.00 Matches: 82
Percent Similarity: 39.25% Conservative: 64
Best Local Similarity: 22.04% Mismatches: 142
Query Match: 5.44% Indels: 84
DB: 6 Gaps: 22

US-10-089-688-1 (1-1053) x US-10-732-923-8708 (1-1359)
QY 19 AAGTCTCAAAATCTCCAGCAGACATTTGTGAAGATCTGAAGAGCAGCATGGCTGTCTG 78
Db 365 LysLysLysGlyLysProGlyLysLeuLysLysGlyIleLeuMetAspSerLysLeu 384
QY 79 GAAAGCAAGACATTTCTGATAAAAGCAAGAAAGGCTACAGAAAGATTTCCAAAAAT 138
Db 385 GlnLysGluThrIleAsp-----AlaGluArgAlaGluLysGluArgArgLysArg 401
QY 139 CTGTTGCCATGAAGAAATTTCTGTATGGC-----ACNAATGAAAGAGCCCTCAGACA 192
Db 402 LeuGluLysLysGlnLysGluPheAsnGlyIleValLeuGluGluGlyGluAspLeuThr 421
QY 193 GAAGCAGTAGCTCAACTGCTCAGAACTCTATAATAGTGGCTCCTTAGCACCCTGGTA 252
Db 422 GluMetLeuThrGlyThrSerSerGlnArgLysLeuLysSerValVal-----LeuAsp 439
QY 253 GCTGATTACAGCTCATTTGAGCGGCAAAAGACGCTGGCTCAAAATTTTCAACAAT 312
Db 440 ProAspSerThrValAspGluGluSerLysLysProVal---GluValHisAsnSer 458
QY 313 ATTCTCAGACAGACAAATTTGGTACGAGAACTCTACTGTGAA-----TACATCTGCACC 366
Db 459 LeuValArgIleLeuLysProHisGlnAlaHisGlyIleGlnPheMetTyrAspCysAla 478
QY 367 CAACAGAAATATTTGTTTCATGTTATTGAAGGGTATGAATCTCAGAAATAGCTCTAAT 426
Db 479 PheGluSerLeuAspArgLeuAspThrGluGly-----SerGlyGlyIleLeuAlaHis 496
QY 427 TGT---GGAATA-----ATGTTAAGAGAATGCATCAGACAT 459
Db 497 CysMetGlyLeuGlyLysThrLeuGlnValIleThrPheLeuHisThrValLeuMetHis 516

```

```

QY 460 GAACCACTTCGAAA-----ATCATT 480
Db 517 GluLysIleGlyLysCysLysArgValLeuValValProLysAsnValIleIle 536
QY 481 TTGTGTCGGAACAGATTTTATGATTTCTTC-----AGATATATGCGAAATGTCAACATTT 534
Db 537 AsnTrpPheLysGluPheGlnLysTrpLeuValAspAsnAspGluGluLeuAspThrIle 556
QY 535 GACATAGCTTCA---GATGCATTTGCCCATTCAGATTTTCTTACAGACATTAATTTG 591
Db 557 AspValAsnGluLeuAspSerTyrLysThrIleGluAsp-----ArgArgAla 573
QY 592 CTCAGTGCAGAAATTTTGGACAGCATTTATGATAGATTTTTCAGTGAATATGAGAGTTA 651
Db 574 Leu-----LysAla 576
QY 652 CTTTATTTCAGAAATTTATGTGCAAAAAGACAGACTCACTGAAGCTTCTCGGTGAACACTA 711
Db 577 TrpHisSer-----SerLysThrProSerValMetIleIleGlyTyrAspLeu 592
QY 712 CTAGATAGACAACTTCAATTTATGACAAATATACATCAGTAAACCTGAGAACCTCAAA 771
Db 593 -----PheArgIleLeuThr---ValGluAspAspProLysLysLysLys 606
QY 772 TTAATCATGAACCTGCTGCGAGACAAAGTCCCAACATCCAGTTTCAG---GCCTTTCAC 828
Db 607 -----ProLysAsnArgAsnArgLeuGluLysAlaLysGlu 619
QY 829 GTTTTAAAGTGTTGTAGCCAATCTTAACAGACGAGCCCATCTAGACATCTCTC--- 885
Db 620 AspPheArgLysTyrLeuGlnAsnProGlyProAspMetValValCysAspGluAlaHis 639
QY 886 ---CTCAAGAAC---CAGGCCAACTCATAGATTTCTCAGCAAGTTTTCAGAACGACAGG 939
Db 640 LysLeuLysAsnAspAspSerAlaLeuSerLysCysMetValLysIleLeuThrLysArg 659
QY 940 ACGGATTGTATGACGAGTCCGTACCGACGACGAAATTCGCGGTGCTGATTTACGC----- 993
Db 660 ArgIleCysLeuThrGlyThr---ProLeuGlnAsnAsnLeuMetGluTyrHisCysMet 678
QY 994 -----GTTAAACCGCGAGCGCTGGGATCAGG 1020
Db 679 ValAsnPheValLysProGlyLeuLeuGlyThrLys 690

RESULT 15
US-10-805-684-90
; Sequence 90, Application US/10805684
; GENERAL INFORMATION:
; APPLICANT: SAKAMOTO, TAKESHI
; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 58748(70342)
; CURRENT APPLICATION NUMBER: US/10805,684
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/455,766
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/459,936
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,103
; PRIOR FILING DATE: 2003-04-02
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 90
; LENGTH: 1960
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-805-684-90

Alignment Scores:
Pred. No.: 0.145 Length: 1960
Score: 97.50 Matches: 76
Percent Similarity: 36.97% Conservative: 63
Best Local Similarity: 20.21% Mismatches: 166

```

```

Query Match:      5.41%      Indels:      71
DB:              6          Gaps:       14

US-10-089-688-1 (1-1053) x US-10-805-684-90 (1-1960)

QY 25 CACAAATCTCCAGCAGCAGCATTGTGAAGATCTGAAGCAGAGCATGGCTGTCTTCTGCAAAAG 84
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1363 HisAlaGlnValThrAspMetCysLysLysMetGluAspGlyValGlyCysLeuGluThr 1382

QY 85 CAAGACATTTCTGATAAAACGAAAGAGCTACAGAAGAGTTTCCAAAAATCTCGTT 144
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1383 AlaGluGluAlaLysArgLeuGlnLysAspGluGlyLeuSerGlnArgLeu--- 1401

QY 145 GCCATGAAGAATAATCTGTATGGCACAAATGAAGAGCCTCAGACAGACAGTAGCT 204
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1402 GluGluLysValAlaAlaTyrAspLysLeuGluLys-----ThrLysThr 1416

QY 205 CAACCTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCTGGTAGCTGATTACAG 264
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1417 ArgLeuGlnGlnGluLeu-----AspAspLeu 1425

QY 265 CTCATTGACTTTGAGGCAAAAGACGCTGGCTCAAAATTTTCAACAATATTCTCAGAAGA 324
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1426 LeuValAspLeuAspHisGlnArg-----GlnSerValSerAsnLeuGluLysLys 1442

QY 325 CAA-----ATTGTCAGAGAACTCTACTGTTGATACATCTGC 363
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1443 GlnLysLysPheAspGlnLeuLeuAlaGluLysThrIleSerAlaLysTyrAlaGlu 1462

QY 364 ACCCAACAGAAATATTTCTTCATGTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTA 423
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1463 GluArgAspArg-----AlaGluAlaGluAlaArgGluLysGluThrLysAlaLeu 1479

QY 424 AATTGTGGAATAATGTTAAGAAATGTCATCAGACATGAACCACTTGCAAAATCATTTTG 483
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1480 SerLeuAlaArgAlaLeuGluGluAlaMet-----GluGlnLysAlaGluLeuGluArg 1497

QY 484 TGGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACATTTTGACATAGCT 543
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1498 LeuAsnLysGlnPheArgThrGluMetGluAspLeuMetSerSerLysAspAspValGly 1517

QY 544 TCAGATGCATTTGCCACATTCAGAGTATTCTTACAGACATAAATG---CTCAGTGCA 600
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1518 LysSerVal-----HisGluLeuGluLysSerLys 1527

QY 601 GAATTTTGGACACAGCATTTATGATAGATTTTTCAGTGAATATGAGAACTTACTTCATCTCA 660
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1528 ArgAlaLeuGluGlnGlnValGluGluMetLysThrGlnLeuGluGlu---LeuGluAsp 1546

QY 661 GAAATATTATGTACAAAAGACAGTCACCTGAAGCTTCTCGGTGAACCTACTACTAGATAGA 720
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1547 GluLeuGlnAlaThrGluAspAlaLysLeuArgLeuGluValAsnLeuGlnAlaMetLys 1566

QY 721 CACAACTTC-----ACAATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTA 774
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1567 AlaGlnPheGluArgAspLeuGlnGlyArgAspGluGlnSerGluGluLysLysGln 1586

QY 775 ATGATGAACCTGCTCGAGACAAAAGTCGCAACATCCAGTTTGAG----- 819
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1587 LeuValArgGlnValArgGluMetGluAlaGluLeuGluAspGluArgLysGlnArgSer 1606

QY 820 ---GCCTTTCAGTTTAAAGTGTGTTAGTCCATCCTAACAGACCGCCCATCCTA 876
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1607 MetAlaMetAlaAlaArgLysLysLeuGluMetAspLeuLysAspLeuGluAlaHisIle 1626

QY 877 GACATCTCTCCTCAGAACAGCCCAAACTCATAGATGTTCTCAGCAAGTTTTCAGAACGAC 936
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1627 AspThrAlaAsnLysAsnArgGluGluAlaIleLysGlnLeuArgLysLeuGlnAlaGln 1646

QY 937 AGGACGGATTTGATGACGAGTTTCCGTACCGACGAGAAATTCGGGGTCGATTTA----- 990
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1647 MetLysAspCysMetArgGluLeuAspAspThrArgAlaSerArgGluGluIleLeuAla 1666

```

Search completed: November 10, 2004, 21:15:05
Job time : 45.5 secs

```

QY 991 CGCGTTAAACCGCGACGCGTGGATCAGCGATTG----- 1026
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1667 GlnAlaLysGluAsnGluLysLysLeuLysSerMetGluAlaGluMetIleGlnLeuGln 1686

QY 1027 -----AAGAGACCAGCTCAGCAAGAA 1047
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1687 GluGluLeuAlaAlaAlaGluArgAlaLysArgGlnAlaGlnGlnGlu 1702

```